The role of noise in optimisation and diffusion limited aggregation

by

Neill E. Bowler

Thesis

Submitted to the University of Warwick

for the degree of

Doctor of Philosophy

University of Warwick, Department of Physics

September 2001

To the glory of God

Contents

List of Tables

List of Figures

Acknowledgments

I would like to thank Robin Ball for his patience in directing me through this small area of physics.

I acknowledge Len Sander and Ellak Somfai for collaboration in the work of chapter 5. Their input in this area was most fruitful.

I would like to thank Steve Goodwin, Peter King and Bob Smith for their help with the work of chapter 4.

I wish to express thanks to everyone who has helped me in my studies. George Rowlands whose assistance in all things mathematical has been vital, particularly for input in chapter 2. Roddy Vann and Jim Hague for advice on computing matters. Mark Barrow for his assistance in figure preparation and many other areas.

Thank you to BP and the EPSRC for financial support during my studentship.

Finally, I want to thank my loving wife, Su. Her love and patience have been, at times, the only thing that has kept me going. "Behind every successful man is a surprised woman." - Maryon Pearson

Intended publications

The following sections of this thesis are intended for publication.

Chapter 2: "Stochastic annealing", Robin C. Ball & Neill E. Bowler, To be submitted to Physical Review Letters, October 2001.

Chapter 3: "Characterization of the probabilistic traveling salesman problem", Neill E. Bowler, Thomas M. Fink & Robin C. Ball, Submitted to Physical Review E, 2001.

Chapter 4: "Oil field optimization under uncertainty", Neill E. Bowler & Robin C. Ball, In preparation, 2001.

Chapter 5: "Off-lattice noise reduction and the ultimate scaling of DLA in two dimensions", Robin C. Ball, Neill E. Bowler, Leonard M. Sander & Ellak Somfai, Submitted to Physical Review Letters, 2001.

Chapter 6: "Noise reduction and the fixed point of DLA in 3 dimensions", Neill E. Bowler & Robin C. Ball, In preparation, 2001.

Abstract

This thesis focuses on the role played by fluctuations in both thermal optimisation techniques and diffusion limited aggregation. The key idea is that by tuning the level of input noise asymptotic results may be attained more rapidly. Stochastic optimisation problems are considered, where the function being optimised cannot be known exactly and may only be estimated. The noise in the estimates is used as the analogue of thermal fluctuations in simulated annealing. This analogy is made exact by use of an acceptance function, and stochastic annealing is seen to be the generalisation of simulated annealing to stochastic optimisation problems. The probabilistic travelling salesman problem (PTSP) is used as a test-bed for stochastic annealing, and significant new results are found. A good characterisation is found for the PTSP and scaling arguments are shown to be accurate for determining the expected length of the pruned and a priori tours, specifically $E(\bar{L}_{\text{pruned}}) = \beta_{\text{pruned}}(p)\sqrt{np}$. An oil field project, as a complex commercial problem, is considered and stochastic annealing is seen to make a large improvement in the expected return of the project.

Noise reduction in diffusion limited aggregation (DLA) is known to be crucial to our understanding. A generalisation of noise reduction off-lattice is introduced, and noise reduction is shown to be a central parameter controlling the growth of DLA. In 2 dimensions, all quantities appear to be influenced by the slowest correction to scaling, with exponent ~ 0.3. In 3 dimensions, some quantities are not affected by the slowest correction to scaling, exponent ~ 0.2. The renormalisation of DLA is considered, and the noise reduction at the fixed point is measured. The noise, given as the relative variability in extremal cluster radius is found to be $\epsilon^*_{2D} \simeq 0.003$ and $\epsilon^*_{3D} \simeq 0.006$ in 2 and 3 dimensions, respectively.

Introduction

Noise is an important aspect of physics, creating imprecision in measurements and obscuring underlying trends. This thesis examines the role of noise in two areas - optimisation and diffusion limited aggregation. Ironically, the latter of these relies on the presence of noise for the appearance of the phenomenon.

In optimisation, noise is an important feature to ensure that an optimisation finds the global optimum. If we are on a rugged terrain, moving only downhill, we may become stuck in a "local" minimum, unable to find the true lowest point of the landscape. However, if we sometimes move uphill, then we may avoid getting stuck and find "global" minimum. This acceptance of uphill moves manifests itself as thermal fluctuations in simulated annealing.

Diffusion limited aggregation (DLA) is a model of Laplacian growth. Growth of smooth objects in a Laplacian field is unstable and the fractal patterns of DLA are the result of this instability. The standard implementation of DLA is noisy, and a way of reducing the influence of noise has been central to the understanding of its eventual behaviour in lattice-based models. What happens when there is no underlying lattice has remained controversial, in large part because a suitable noise reduction technique have not been available.

Chapter 1 gives an introduction to optimisation and stochastic optimisation. We consider the use of heuristics (approximate methods) to solve hard optimisation problems. Stochastic optimisation problems are introduced, in which the exact value of the function being optimised is not know exactly and must be estimated, and previous solution methods are reviewed.

Chapter 2 is centred around a particular stochastic optimisation technique, called stochastic annealing. The noise in the evaluations of the function being opti-

mised are considered as the exact analogue of the thermal fluctuations in simulated annealing. This is achieved by choosing an acceptance function, and the appropriate acceptance function is found for a number of distributions of the estimates. Hence, stochastic annealing is taken to be the generalisation of simulated annealing to stochastic optimisation problems.

Chapters 3 & 4 are concerned with two applications of stochastic annealing. The probabilistic travelling salesman problem is a model of academic value which is known to be hard, for which a good characterisation is provided. Oil field projects are complex projects which provide a demonstration of the utility of stochastic annealing in a commercial setting.

Chapters 5 & 6 concentrate on DLA and the importance of noise reduction. A new and efficient noise reduction technique is described for off-lattice DLA which may be applied in any dimension. Asymptotic values and correction to scaling exponents for isotropic DLA are found. The renormalisation of DLA is considered and the fixed point noise reduction is found, which leads to consideration of an effective noise reduction. The first ever implementation of noise reduction off-lattice for DLA in 3 dimensions is achieved.

Chapter 1

Optimisation and stochastic optimisation

Research! A mere excuse for idleness; it has never achieved, and will never achieve any results of the slightest value.
– Benjamin Jowett

Optimisation may be described as finding the extremal value of a function. If one considers a simple function, say $E(y) = y^2$, the minimum of such a function may be found very quickly. However, the more complex the function the more difficult this analysis becomes, particularly in the case of NP-complete problems. Furthermore, in some cases one may not be able to exactly evaluate the function E.

In this chapter the topic of NP-completeness is introduced as are some techniques which are designed to solve NP-complete problems. Stochastic optimisation is then described, and the specific challenge posed by stochastic optimisation problems is considered.

1.1 NP-completeness

1.1.1 What it is

We may consider partitioning optimisation problems into two sets, which we will describe as "easy" and "hard". Easy problems are ones for which we can find the solution in a time polynomial in the size of the problem. Hard problems are ones for which we cannot guarantee finding the solution in a time polynomial in the size

of the problem (i.e. the time required to find the solution grows exponentially or faster with the problem size).

NP-complete is an important class of problems which are *believed* to be hard (see [21], chapter 9). Many combinatorial optimisation problems are NP-complete, such as the travelling salesman problem (TSP) and prime factorisation. Although these problems are believe to be hard, this has never been proved.

NP-complete problems are not the hardest problems possible. There are problems whose solution requires an exponential time to write. However, the importance of the class of NP-complete problems, since they are simple to state but difficult to solve, should not be neglected.

1.1.2 Why it's important - need for heuristics

The class of hard problems is important, since it defines when it is no longer feasible to search for exact solutions. Although it is possible to find an exact solution to a hard problem, the size of the problem that may be solved will always be very limited. Hence a popular approach is to set a fixed time for an algorithm to run, and take the best solution found within that time. This approach is only sensible with problems for which a near-optimal solution is useful. Some problems, such as prime factorisation, require an exact solution, and it is this requirement that makes internet security so difficult to break.

In the next two sections we consider some heuristics, concentrating on simulated annealing, as this has been the subject of the work presented in the following chapters.

1.2 Simulated Annealing

Annealing is a process by which a material (such as glass) is heated to a moderate temperature to remove defects and soften the material for working. Kirkpatrick et al. [53] used this process as inspiration for an optimisation technique, simulated annealing. One considers the optimisation problem as a thermal system in equilibrium at a given temperature. If this system is allowed to cool slowly, then at zero temperature it will occupy the ground state, and the optimisation problem will be

solved. Simulated annealing is a powerful, general optimisation technique that has found application in many problems.

1.2.1 Thermodynamics & Markov Chains

A system held at a fixed temperature T will occupy a state, ν, with energy $E(\nu)$ according to the probability

$$P(\nu) \propto e^{-\beta E(\nu)} \tag{1.1}$$

where $\beta = \frac{1}{k_B T}$ is the inverse temperature, and k_B is Boltzmann's constant. In the following we chose $k_B = 1$ for convenience. We will attempt to simulate a system for which the states are chosen according to this rule.

When using Monte Carlo methods, we want to efficiently sample the distribution of states of the system so that we may estimate average quantities, such as the average energy of the system. It would not be efficient to attempt to implement eq. 1.1 directly, instead it is convenient to use a Markov process. The initial state of the system is chosen randomly and the system is then allowed to make a transition (often called a move) from one state to another, with a probability which depends only on the states involved in the move.

For each move that we propose, we choose to accept that move with a given probability ($A(\mu \to \nu)$ for a move from state μ to state ν). We impose the correct equilibrium distribution via the condition of detailed balance, which is

$$P(\mu) A(\mu \to \nu) = P(\nu) A(\nu \to \mu). \tag{1.2}$$

Eq. 1.1 is then satisfied provided

$$\frac{A(\mu \to \nu)}{A(\nu \to \mu)} = e^{-\beta \Delta E} \tag{1.3}$$

where ΔE is the energy difference between the states ν and μ.

1.2.2 Metropolis and Glauber algorithms

Whilst eq. 1.3 ensures thermal equilibrium at inverse temperature β, it does not fully determine the form of A. This means that there are a number of different implementations of simulated annealing. A popular choice (which ensures A is as

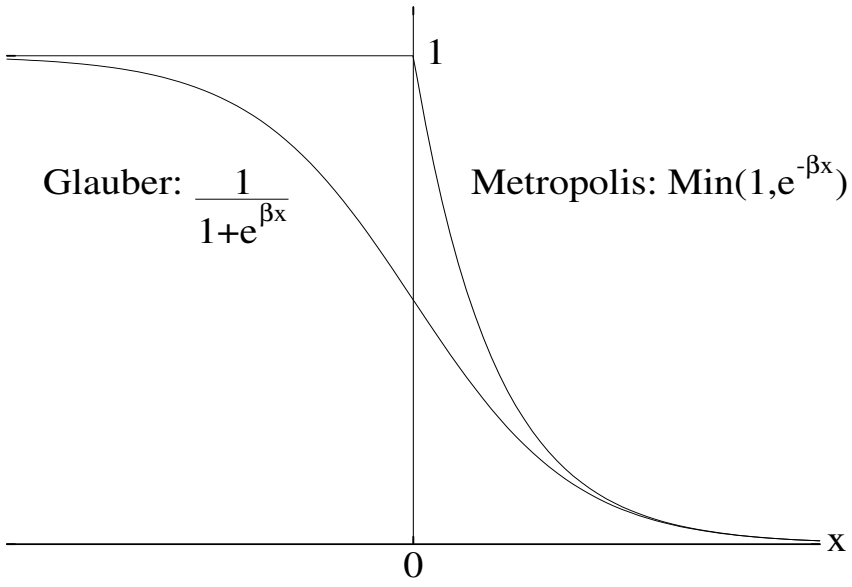

Figure 1.1: The Glauber and Metropolis acceptance functions. If we are simulating a thermal system at inverse temperature β and measure the effect of a move to be ΔE, then we may choose to accept moves with a probability given by the above functions

large as possible) is to choose $A(\Delta E) = 1$ for $\Delta E < -C$, with C as small as possible. This is the Metropolis algorithm, and is defined by

$$A(\Delta E) = \begin{cases} 1 & \text{for } \Delta E < 0 \\ e^{-\beta \Delta E} & \text{for } \Delta E > 0 \end{cases} \tag{1.4}$$

since eq. 1.3 ensures that $C = 0$. A different choice of A is to insist that $A(\Delta E) + A(-\Delta E) = 1$. This leads to the Glauber acceptance rule, given by

$$A(\Delta E) = \frac{1}{1 + e^{\beta \Delta E}} \tag{1.5}$$

which has the benefit of being analytic. Both acceptance rules are illustrated in figure 1.1.

1.2.3 Temperature schedules

If we apply the simulated annealing algorithm, we hope to simulate a system in thermal equilibrium at some temperature. We consider starting the system at a high temperature, and gradually cooling the system to zero temperature. At zero

temperature, the system will occupy the ground state (or equivalently have solved the optimisation problem) provided we have remained in equilibrium. However, to stay in equilibrium down to zero temperature requires an infinite amount of time. In fact, Geman & Geman [38] showed that if one varies the temperature according to

$$T = \frac{F}{\ln t} \tag{1.6}$$

where F is a constant which must be chosen to be sufficiently large, then the system will occupy the ground state with probability 1, as the time through the simulation, $t \to \infty$.

Unfortunately, this is not much use, since we never have an infinite amount of time. There have been a number of suggestions for good schedules when time is limited [99, 85, 20]. Most notable is the schedule which maintains a constant thermodynamic speed [86, 95]. The principle is a simple - if one cannot stay in equilibrium, one should attempt to stay at a fixed (thermodynamic) distance from equilibrium. This leads to the suggestion to use

$$\frac{dT}{dt} = -\frac{\vartheta T}{\epsilon\sqrt{C}} \tag{1.7}$$

where ϵ and C are the relaxation time and heat capacity of the system, respectively. ϑ is the thermodynamic speed, a constant one chooses to determine the overall speed of the cooling. In principle this is the optimal schedule to use. However, it requires that we know the relaxation time of the system, ϵ, which is not simple to measure. Rees & Ball [91] also used a schedule which utilised the relaxation time.

In this thesis much more simple schedules have been used, based more on intuition and what appears to work than physical principles.

1.3 Other heuristics

Simulated annealing is not the only heuristic algorithm, and is by no means the simplest. Other algorithms have proved useful in different situations, and the choice of which heuristic to use is an important one. To give a flavour of the diversity, a few other algorithms are discussed.

1.3.1 Wolff algorithm

The Wolff algorithm [116, 108] is a variant of simulated annealing, designed for efficient simulation of the Ising model. The Ising model is a microscopic model of a ferromagnet where each site of a lattice possesses a spin which can be either up or down. Adjacent spins prefer to be parallel, with a fixed energy penalty for being anti-parallel. At low temperatures, simulated annealing is inefficient at simulating the Ising model. If we consider a move to flip a single spin, this will most likely be rejected, since most spins will be surrounded by many other spins pointing in the same direction.

The Wolff algorithm avoids this low temperature problem by flipping clusters of spins at the same time. The algorithm maintains a thermal selection rule, with each move being selected in a manner so that every proposed move is accepted.

Simulated annealing is inefficient at low temperatures, staying in one configuration for a long time. This is similar to the concept of a local minimum in optimisation. Imagine that we are seeking to minimise $E(y)$ shown in figure 1.3.1. The system starts in some initial configuration, y_0, and after a number of moves the system is in state y_1. This is a local minimum, and the system will have to accept a number of unfavourable moves to escape this minimum. If the temperature is low, it will take a long time to escape this configuration. In the Ising model, there is no difference between the system having all spins up or all spins down. At low temperatures, simulated annealing will not be able to sample both these states since they are separated by an energy barrier. It is for this reason that the temperature schedule in simulated annealing is so important.

1.3.2 Tabu search

A "tabu search" [39, 40] is an attempt to avoid the problem of getting stuck in local minima. One constructs a list of the previous n configurations of the system which are taboo, and may not be visited in the next move. This way we may be forced to accept unfavourable moves, since all favourable moves may be taboo.

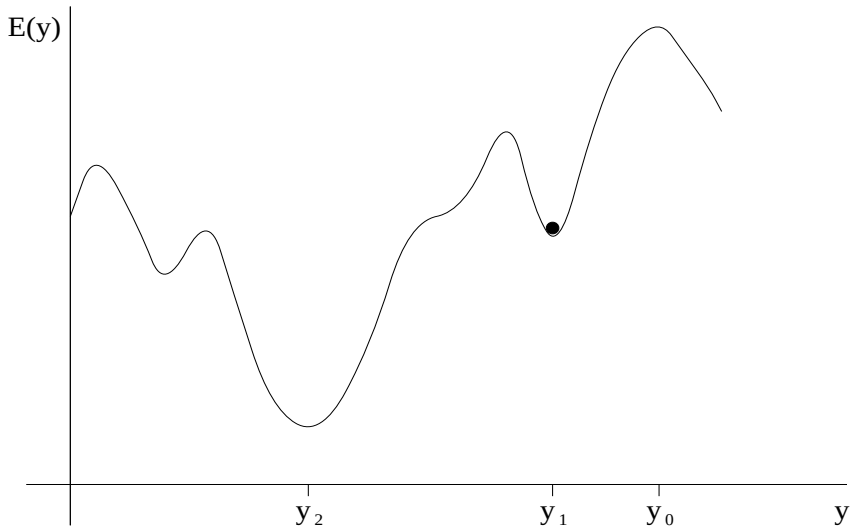

Figure 1.2: Getting stuck in a local minimum. We consider optimising $E(y)$ by accepting only "downhill" moves (i.e. ones which decrease the value of E). If the system starts at y_0 then it will get stuck in local minimum, y_1, and will not be able to find the global minimum, y_2.

1.3.3 Local search

When implementing simulated annealing, we propose a move from the systems current configuration to a neighbouring configuration. We are free to chose which type of moves (called the move-set) to consider, and thus we may choose the structure of the local neighbourhood. Many local search heuristics are based on the hope that we can choose the move-set so as to ensure that the function has a single, global minimum, and not many local minima. If this can be achieved, then a simple algorithm which accepts only favourable moves will solve the problem.

This approach was very popular when computer time was restricted and the use of global search heuristics was not feasible. Many have applied local search heuristics with reasonable success [63, 64, 22]. The choice of move-set is also an important aspect when using simulated annealing, as this helps to make the algorithm more efficient.

1.4 Stochastic Optimisation

Previously, we have assumed that the function we wish to optimise, $E(y)$, is known exactly. A more general situation, known as stochastic optimisation, is where $E(y)$ is not known, but may be estimated. There is some underlying function that we wish to optimise, and our estimates of this are distributed around this value, $E(y)$ being the mean value of the estimates. If we assume each estimate to be statistically independent, then we may consider each move independently, and also average repeated estimates to gain more precise knowledge of $E(y)$.

Each estimate can be thought of as giving the underlying value subject to some noise. The noise in the estimates needs to be kept small if low temperatures are to be simulated, while at high temperatures low quality estimates may be used. The difficulty in the problem is now transferred from the complexity of E to also considering how precise an estimate is needed, so as to avoid spending unnecessary effort.

1.4.1 Previous optimisation methods

Stochastic optimisation was borne out of an idea by Robbins & Monro [92]. They considered solving the problem of finding

$$\Gamma(y) = \alpha \tag{1.8}$$

where Γ is some monotonic function of y, the solution to be found, and α is a parameter. Γ is not found directly, but can only be estimated. Their technique to solve this problem is called stochastic approximation. A number of variants of this scheme have since been developed [12, 56, 59], including a generalisation to be able to tackle stochastic optimisation problems [55].

Stochastic optimisation is a more general situation, where we wish to find the minimum of some function, but where this function may have many minima. Just as we may consider ordinary optimisation methods to fall into two categories - exact methods and heuristics - stochastic optimisation methods can also be classed as developments from these two categories of method. Two of the developments of exact methods are called stochastic linear programming [10, 23, 45] and stochastic dynamic programming [93].

A number of heuristics exist to tackle stochastic optimization problems [2, 29, 41, 117, 118], with varying advantages and disadvantages. For example consider the stochastic comparison algorithm [41], which is a development of an algorithm by Yan and Mukai [118]. With this algorithm, one makes a number of estimates of the energy of the current state, $X_t^{i=1,r}$, and of the new state $X_{t+1}^{i=1,r}$. If any of the estimates of the energy of the new state are greater than the corresponding estimate of the energy of the current state, then the current state is kept and the proposed move rejected. It is quite clear that this is an efficient method for high temperatures, but will reject many favourable moves at low temperature.

Many of the heuristics are developments from simulated annealing [1, 18, 43, 52]. Simulated annealing [43] has been shown [42] to solve stochastic optimization problems with probability 1 provided the noise in the estimates at time step t is kept below $O(t^{-\gamma})$ where $\gamma > 1$.

Modified simulated annealing algorithms [1, 18, 52] take into account the width of the distribution of the estimates to make sure that the estimates of the energy are significantly different before accepting or rejecting a move. Some of these methods have been shown to converge to the global minimum of the function being optimised under given conditions [37, 52].

Stochastic annealing [34] is an extension of simulated annealing to stochastic optimisation problems. It differs from the above approaches in two key areas. Firstly, the general concept of the other approaches is one of minimising the noise present in the estimates. It is entirely natural that the noise present in the energy evaluations should be used as a positive effect, since one may think of simulated annealing as a noisy quench. Secondly, this approach can be modified to give exact simulation of a thermal system. Although this is not specifically ruled out by the other approaches, no attempt has been made to satisfy this condition.

1.4.2 Acceptance probabilities in stochastic systems

In the stochastic optimisation context the probability $P_{accept}(\Delta E)$ of accepting a move whose energy change is ΔE is not the same as the probability $A(x)$ of accepting moves whose energy change appears to be x. These are two distinct quantities, since we no longer have direct access to ΔE. Let $f(x|\Delta E)$ be the probability density

of estimating that the energy change is x when its true value is ΔE. Then the probability of accepting a move whose true energy change is ΔE is given by

$$P_{accept}(\Delta E) = \int_{-\infty}^{\infty} f(x \mid \Delta E)\, A(x)\, dx. \tag{1.9}$$

This formula is applicable to stochastic optimisation in general, but so far has only been considered by Fink & Ball [34]. If we can choose $A(x)$ such that P_{accept} follows eq. 1.3, then we will be implementing a generalised version of simulated annealing, an issue which we tackle in the next chapter

Chapter 2

Stochastic annealing

If we knew what it was we were doing, it would not be called research, would it? – Albert Einstein

Stochastic optimistation problems are ones in which we wish to perform an optimisation where exact knowledge of the function to be optimised is not available. It is assumed that increasingly precise estimates of the function are available, but that these estimates are increasingly expensive. The task is to perform an optimisation that is not prohibitively expensive.

The approach of stochastic annealing is a generalisation simulated annealing to stochastic optimisation problems. The original Fink & Ball [34] algorithm only gives an approximation to true thermal sampling. This technique may be adapted to give exactly thermal sampling, and this is demonstrated in a range of cases.

2.1 Stochastic annealing - Fink and Ball

We have already seen (section 1.4.2) that the acceptance probability of a move whose underlying energy change is ΔE is

$$P_{accept}(\Delta E) = \int_{-\infty}^{\infty} A(x) f(x|\Delta E)\, dx. \tag{2.1}$$

$A(x)$ is the probability that we choose to accept a move whose energy change is estimated to be x and $f(x|\Delta E)$ is the probability density of estimating the energy change of a move to be x when its true value is ΔE. Fink and Ball [34] chose $A(x) = 1$ for $x < 0$ and $A(x) = 0$ otherwise. This choice stems from the concept of a probabilistic quench - that we accept all moves which *appear* to be favourable.

If we have more precise estimates of the energy change, then we will accept fewer unfavourable moves - equivalent to lowering the temperature. Thus, changing the precision of our estimates of ΔE allows us to perform an annealing.

Fink and Ball[34] considered the particular case of f being a Gaussian disribution

$$f(x|\Delta E) = \frac{e^{\frac{-(x-\Delta E)^2}{2\sigma^2}}}{\sqrt{2\pi}\sigma}. \tag{2.2}$$

Substitution of this into eq. 2.1 with the choice of A above gives

$$P_{accept}(\Delta E) = \frac{1}{2}\left[1 - \mathrm{erf}\left(\frac{\Delta E}{\sqrt{2}\sigma}\right)\right]. \tag{2.3}$$

This is sketched in fig. 2.1, which illustrates the principle behind stochastic annealing. To assure thermal equilibrium, the acceptance probabilities must obey (eq. 1.3)

$$\frac{P_{accept}(\Delta E)}{P_{accept}(-\Delta E)} = e^{-\beta \Delta E} \tag{2.4}$$

where $\beta = \frac{1}{k_B T}$ is the inverse temperature and ΔE is the energy difference between the states in the move. Using the Fink & Ball choice of $A(x)$ does not give an acceptance probability which satisfies eq. 2.4 exactly, but closely approximates it. Using eq. 2.3 and expanding in ΔE we obtain

$$\begin{aligned} ln\left(\frac{P_{accept}(\Delta E)}{P_{accept}(-\Delta E)}\right) &= ln\left(\frac{1-\mathrm{erf}\frac{\Delta E}{\sqrt{2}\sigma}}{1+\mathrm{erf}\frac{\Delta E}{\sqrt{2}\sigma}}\right) \\ &\simeq -\beta_G \Delta E - \tfrac{4-\pi}{48}(\beta_G \Delta E)^3 - \ldots \end{aligned} \tag{2.5}$$

where

$$\beta_G = \frac{\sqrt{8}}{\sqrt{\pi}\sigma} \tag{2.6}$$

identifies the effective inverse temperature. The small coefficient ($\simeq 0.02$) of the cubic term in eq. 2.5 makes this a rather good approximation to true thermal selection.

To control the temperature we need to be able to control σ. This may be achieved by making a number of statistically independent estimates of ΔE from the Gaussian distribution. The mean value of a set of r estimates is Gaussianly distributed around ΔE, with standard deviation

$$\sigma_r = \frac{\sigma}{\sqrt{r}} \tag{2.7}$$

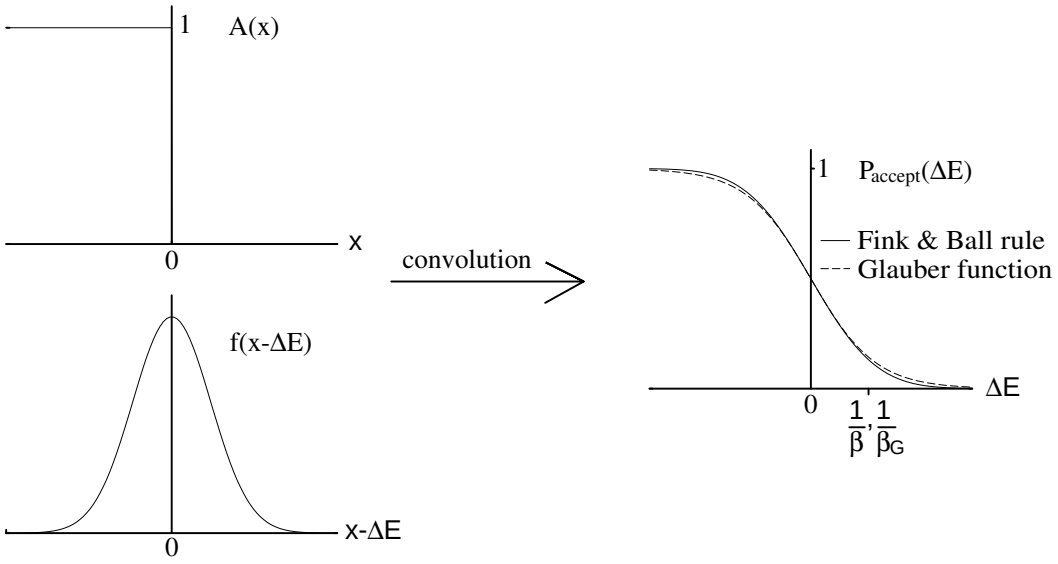

Figure 2.1: The principle behind stochastic annealing. One chooses to accept moves with a probability $A(x)$. Since x is an estimate of ΔE, $P_{accept}(\Delta E)$ is the convolution of $A(x)$ with the error distribution $f(x - \Delta E)$. The Glauber acceptance function is also shown for comparison.

where r is the sample size. If we use this as our estimate of ΔE, the effective inverse temperature becomes

$$\beta_G = \frac{\sqrt{8r}}{\sqrt{\pi}\sigma} \tag{2.8}$$

and is tunable by varying r. This sampling method also has the advantage that for large r the central limit theorem drives the distribution of the average of the estimates towards a Gaussian even if it is not so for $r = 1$. Note that for a smooth annealing schedule one should take care to measure σ and control β_G rather than just r alone. This turns out to be important for the probabilistic travelling salesman problem, see sec. 3.5.

2.2 Naïve Approach

If we make a different choice of $A(x)$ to Fink and Ball [34], we can hope to satisfy the detailed balance condition, eq. 2.4, exactly. We know that for simulated annealing we can make a number of different choices of acceptance function, such as the Glauber or Metropolis acceptance rules (section 1.2.2). If we insist that P_{accept} follows one

of these rules, then we will have satisfied eq. 2.4. For example eq. 2.1 would lead to

$$P_{Glauber}(\Delta E) = \int_{-\infty}^{\infty} A(x)\, f(x|\Delta E)\, dx. \tag{2.9}$$

In the case where $f(x|\Delta E) = f(x - \Delta E)$, this is a de-convolution problem for $A(x)$. Taking Fourier-Laplace transforms of both sides, we have

$$\widetilde{A}(p) = \frac{\widetilde{P}_{Glauber}(p)}{\widetilde{f}(p)} \tag{2.10}$$

where the Fourier-Laplace transform is defined by

$$FLT(f(x)) = \widetilde{f}(p) = \int_{-\infty}^{\infty} e^{-px} f(x)\, dx. \tag{2.11}$$

In the case of the Glauber acceptance rule (eq. 1.5), with f as the Gaussian distribution (eq. 2.2), this becomes

$$\widetilde{A}(p) = \frac{2i\pi}{\beta} \frac{e^{\frac{p^2\sigma^2}{2}}}{e^{\frac{i\pi p}{\beta}} - e^{-\frac{i\pi p}{\beta}}} \tag{2.12}$$

The difficulty is that $A(x)$ must be a probability, and thus lie in the range $[0, 1]$ for all x. There is no obvious way to ensure that this condition is satisfied, since eq. 2.12 fully specifies $A(x)$. We use a less restrictive approach below which allows some freedom in our choice of A, and hence we may choose it as a probability.

2.3 Simulating a thermal system

2.3.1 General case

In the following analysis we will be attempting to choose an $A(x)$ such that P_{accept} satisfies detailed balance, eq. 2.4. We will assume that the function $f(x|\Delta E)$ is known exactly. Combining eq. 2.4 with eq. 2.1 we have

$$\int_{-\infty}^{\infty} A(x)\, f(x|\Delta E)\, dx = e^{-\beta \Delta E} \int_{-\infty}^{\infty} A(x)\, f(x| - \Delta E)\, dx. \tag{2.13}$$

We make the substitutions $f(x|\Delta E) = e^{\frac{\beta}{2}(x-\Delta E)} g(x, \Delta E)$ and $A(x) = e^{-\frac{\beta x}{2}} a(x)$, giving

$$\int_{-\infty}^{\infty} a(x)\ (g(x, \Delta E) - g(x, -\Delta E))\, dx = 0. \tag{2.14}$$

As a solution, we try

$$a(x) = \int_{-\infty}^{\infty} h(\Delta E')\, j(x, \Delta E')\, d\Delta E' \tag{2.15}$$

where h and j are functions to be chosen. Substituting this into equation 2.14 and re-arranging we find

$$\int_{-\infty}^{\infty} h(\Delta E') \left[k(\Delta E, \Delta E') - k(-\Delta E, \Delta E')\right] d\Delta E' = 0 \tag{2.16}$$

where

$$k(\Delta E, \Delta E') = \int_{-\infty}^{\infty} g(x, \Delta E) j(x, \Delta E')\, dx. \tag{2.17}$$

Eq. 2.16 is satisfied when $h(\Delta E') = h(-\Delta E')$ and $k(\Delta E, \Delta E') = k(-\Delta E, -\Delta E')$ are even functions. We have not succeeded in taking this most general case further, the hard part being to ensure that $0 \leq A(x) \leq 1$ for a probability.

2.3.2 Specific case

We now make the transition to a less general case for the error distribution. We consider that $f(x|\Delta E)$ is a function of $x - \Delta E$ alone, meaning that the distribution of error, $x - \Delta E$, is independent of ΔE itself. We choose this in part because if we allow the form of f to depend on ΔE, then there is no reason for not allowing it to depend on some other variable, the case above being essentially identical to the most general case possible.

For this case, we choose that $j(x, \Delta E') = g(x - \Delta E')$ and substitute into eq. 2.17 to get

$$k(\Delta E, \Delta E') = \int_{-\infty}^{\infty} g(x - \Delta E) g(x - \Delta E')\, dx. \tag{2.18}$$

This is an even function in the sense that if we substitute $y = x + \Delta E + \Delta E'$ into this, we find

$$\begin{aligned} k(\Delta E, \Delta E') &= \int_{-\infty}^{\infty} g(y + \Delta E) g(y + \Delta E')\, dy & (2.19) \\ &= k(-\Delta E, -\Delta E'). & (2.20) \end{aligned}$$

Therfore,

$$a(x) = \int_{-\infty}^{\infty} h(\Delta E')\, g(x - \Delta E')\, d\Delta E' \tag{2.21}$$

is a solution to eq. 2.14 provided h is even. Taking Fourier-Laplace transforms of this we have

$$\widetilde{a}(p) = \widetilde{h}(p)\widetilde{g}(p). \tag{2.22}$$

We still have some choice in $\widetilde{h}$, since the only restriction we have so far specified is that it is even ($\widetilde{h}$ must be even, since h is even). Given it is desirable to reject as few moves as possible, we choose $A(x)$ $(= e^{-\frac{\beta x}{2}} a(x))$ to be identically 1 below some threshold for x. Hence we define

$$\widetilde{a}(p) = \int_{-\infty}^{S_o} e^{\frac{\beta x}{2}} e^{-px}\, dx + \int_{S_o}^{\infty} \alpha(x) e^{-px}\, dx \tag{2.23}$$

where $\alpha(x) = A(x) e^{\frac{\beta x}{2}}$ for $x \geq S_o$ and $\alpha(x) = 0$ for $x < S_o$. Thus we have

$$\widetilde{\alpha}(p) = \widetilde{h}(p)\widetilde{g}(p) - \frac{e^{\left(\frac{\beta}{2}-p\right)S_o}}{\frac{\beta}{2}-p} \text{ for } Re(p) < \frac{\beta}{2}. \tag{2.24}$$

We need to ensure that $P_{accept}(\Delta E)$ $\left(= \int_{-\infty}^{\infty} A(x) f(x - \Delta E)\, dx\right)$ corresponds to a probability. This means that $P_{accept}(\Delta E) \leq 1$ for all ΔE and since we are insisting on eq. 2.4 then

$$P_{accept}(\Delta E) \leq e^{-\beta \Delta E} \tag{2.25}$$

for all ΔE. Hence we have

$$\int_{-\infty}^{\infty} A(x) f(x - \Delta E)\, dx \leq e^{-\beta \Delta E}. \tag{2.26}$$

This leads to restrictions on $A(x)$ and $f(x - \Delta E)$, since both are positive. Since we are choosing $A(x) = 1$ for $x < S_o$ we have

$$\int_{-\infty}^{S_o} f(x - \Delta E) dx \leq e^{-\beta \Delta E} \tag{2.27}$$

which means (quite reasonably) that f must fall off at least as fast as the distribution of states we are attempting to simulate. From eq. 2.26 we have

$$\int_{S_o - \Delta E}^{\infty} A(y + \Delta E) f(y)\, dy \leq e^{-\beta \Delta E}. \tag{2.28}$$

Since $A(x) \to 0$ as $x \to \infty$, it is reasonable to assume that A is concave; that is

$$A(y + \Delta E) \geq A(\Delta E) + y A'(\Delta E). \tag{2.29}$$

Combining this with eq. 2.28 gives

$$A(\Delta E) \leq e^{-\beta \Delta E} \quad \text{for } \Delta E \gg 0. \tag{2.30}$$

To calculate the limit this places on $\widetilde{\alpha}(p)$, we write

$$\widetilde{\alpha}(p) \leq \int_{S_o}^{\infty} \left| A(x) e^{\frac{\beta x}{2}} e^{-px} \right| dx \tag{2.31}$$

which follows from eq. 2.23. From this, we find

$$\widetilde{\alpha}(p) \leq \int_{S_o}^{\infty} Be^{-\beta x} e^{\left(\frac{\beta}{2} - Re(p)\right)x} dx \tag{2.32}$$

where B is some constant, which leads to

$$\widetilde{\alpha}(p) \leq B \frac{e^{-\left(\frac{\beta}{2} + Re(p)\right)S_o}}{\frac{\beta}{2} + Re(p)} \tag{2.33}$$

Hence, $\widetilde{\alpha}(p)$ is bounded for all $Re(p) > -\frac{\beta}{2}$. Since $\widetilde{g}(p)$ is not guaranteed to be bounded for all p, h must be chosen to ensure that this condition on $\widetilde{\alpha}$ is met. Using the Wiener-Hopf method [79] we may define

$$\widetilde{g}(p) = \widetilde{g_L}(p)\widetilde{g_R}(p) \tag{2.34}$$

where $\widetilde{g_L}(p)$ is non-zero and bounded for $p < 0$, and similarly for $\widetilde{g_R}(p)$ for $p > 0$.

Since $\widetilde{\alpha}$ is bounded for $Re(p) > -\frac{\beta}{2}$ we need to take account of the pole contained explicitly in eq. 2.24. We chose $\widetilde{h}(p)$ to contain a pole of opposite residue at $p = \frac{\beta}{2}$, we can cancel the effect of the pole in eq. 2.24. If $\widetilde{h}(p) \propto \frac{1}{\widetilde{g_L}(p)}$ we ensure that $\widetilde{h}(p)\widetilde{g}(p)$ is bounded for $Re(p) > -\frac{\beta}{2}$. Since $\widetilde{h}(p)$ must be even, we have

$$\widetilde{h}(p) = \frac{\beta}{\left(\left(\frac{\beta}{2}\right)^2 - p^2\right)} \frac{\widetilde{g_L}\left(-\frac{\beta}{2}\right)}{\widetilde{g_R}\left(\frac{\beta}{2}\right)\widetilde{g_L}(p)\widetilde{g_L}(-p)} \tag{2.35}$$

which gives

$$\widetilde{\alpha}(p) = -\frac{e^{\left(\frac{\beta}{2} - p\right)S_o}}{\frac{\beta}{2} - p} + \frac{\beta}{\left(\left(\frac{\beta}{2}\right)^2 - p^2\right)} \frac{\widetilde{g_R}(p)\widetilde{g_L}\left(-\frac{\beta}{2}\right)}{\widetilde{g_R}\left(\frac{\beta}{2}\right)\widetilde{g_L}(-p)}. \tag{2.36}$$

Solution of this equation leads to the optimal acceptance rule for $f(x)$. It may seem that there is still some freedom in the choice of $\widetilde{\alpha}(p)$. This is not the case, since the condition on $\widetilde{\alpha}(p)$ (eq. 2.33) will define S_o.

We now go on to discuss some particular functions f.

2.4 Results

A summary of the results obtained using this method when applied to various error distributions, f, is shown in table 2.1. In section 2.4.1 we consider the simple case

Name $f(x-\Delta E)$	$A(x)$
Top Hat $\frac{1}{C}$ for $\lvert x-\Delta E\rvert < \frac{C}{2}$ 0 otherwise	1 for $x < -\frac{C}{2}$ $\frac{e^{-\beta\left(x-\frac{C}{2}\right)}(\beta(x+C)+1)-1}{e^{\beta C}-1}$ for $-\frac{C}{2} < x < \frac{C}{2}$ $\frac{\beta C e^{-\beta\left(x-\frac{C}{2}\right)}}{e^{\beta C}-1}$ for $x > \frac{C}{2}$
Convolved Exponentials $\frac{n\alpha}{2}e^{-\alpha\lvert x_1\rvert} \otimes \frac{n\alpha}{2}e^{-\alpha\lvert x_2\rvert} \otimes \ldots \otimes \frac{n\alpha}{2}e^{-\alpha\lvert x_n\rvert}$	1 for $x < 0$ $\left(\frac{\beta+\alpha n}{\alpha n}\right)^n \sum_{i=0}^{n} \frac{(-1)^i n!\beta^i}{i!(n-i)!}$ $\left[e^{-\beta x}\left(\frac{1}{\alpha n}\right)^i - \left(\frac{1}{\alpha n+\beta}\right)^i + \left(\sum_{j=1}^{i} \frac{x^{j-1}e^{-(\beta+\alpha)x}}{(j-1)!}\left(-\frac{\beta}{\alpha n(\alpha n+\beta)}\right)^j\right)\right] + 1$ for $x > 0$

Table 2.1: The optimal choice of the acceptance probability, $A(x)$, for a given error distribution, f, using our method. Some more complicated examples are found in table 2.2

of the exponential distribution, and the Gaussian distribution in section 2.4.2. In section 2.4.3 we consider the "modified" Gaussian distribution, which proves to be much more troublesome. Results for this technique when applied to bimodal distributions are summarised in table 2.2.

2.4.1 Exponential distribution

We define the exponential distribution as

$$f(x-\Delta E) = \frac{\gamma}{2}e^{-\gamma|x-\Delta E|} \tag{2.37}$$

which leads to

$$\widetilde{g}(p) = \frac{\gamma^2}{\left(\gamma-\frac{\beta}{2}-p\right)\left(\gamma+\frac{\beta}{2}+p\right)}. \tag{2.38}$$

We choose

$$\widetilde{g_L}(p) = \frac{\gamma}{\left(\gamma-\frac{\beta}{2}-p\right)} \tag{2.39}$$

$$\widetilde{g_R}(p) = \frac{\gamma}{\left(\gamma+\frac{\beta}{2}+p\right)}. \tag{2.40}$$

This leads to

$$\widetilde{\alpha}(p) = -\frac{e^{\left(\frac{\beta}{2}-p\right)S_o}}{\frac{\beta}{2}-p} + \frac{\beta}{\left(\frac{\beta}{2}\right)^2-p^2}\frac{\gamma+\beta}{\gamma}\frac{\gamma-\frac{\beta}{2}+p}{\gamma+\frac{\beta}{2}+p}. \tag{2.41}$$

Applying eq. 2.33, $S_o = 0$. Re-arrangement gives

$$\widetilde{\alpha}(p) = \frac{\gamma^2 - \beta^2}{\gamma^2\left(p + \frac{\beta}{2}\right)} + \frac{\beta^2}{\gamma^2\left(p + \frac{\beta}{2} + \gamma\right)}. \tag{2.42}$$

The inverse transformation can be performed by inspection to give

$$A(x) = \begin{cases} 1 & \text{for } x < 0 \\ \frac{\gamma^2-\beta^2}{\gamma^2}e^{-\beta x} + \frac{\beta^2}{\gamma^2}e^{-(\gamma+\beta)x} & \text{for } x > 0. \end{cases} \tag{2.43}$$

We note that this acceptance function only remains positive for all x when $\gamma \geq \beta$ as anticipated (eq. 2.27).

2.4.2 Gaussian distribution

We attempt to find a thermal acceptance rule for the Gaussian distribution (eq. 2.2). From this we find

$$\widetilde{g}(p) = e^{\left(p+\frac{\beta}{2}\right)^2\frac{\sigma^2}{2}}. \tag{2.44}$$

This function has an essential singularity at ∞, so $\widetilde{g_L}(p)$ and $\widetilde{g_R}(p)$ cannot strictly be chosen. Ignoring this difficulty, we choose

$$\widetilde{g_L}(p) = \widetilde{g_R}(p) = e^{\left(p+\frac{\beta}{2}\right)^2\frac{\sigma^2}{4}} \tag{2.45}$$

which leads to

$$\widetilde{\alpha}(p) = \frac{e^{\left(\frac{\beta}{2}-p\right)S_o} - e^{\left(p-\frac{\beta}{2}\right)\frac{\beta\sigma^2}{2}}}{p - \frac{\beta}{2}} + \frac{e^{\left(p-\frac{\beta}{2}\right)\frac{\beta\sigma^2}{2}}}{\frac{\beta}{2} + p} \tag{2.46}$$

Since $\widetilde{\alpha}(p)$ must decay as e^{-pS_o} as $p \to \infty$ (eq. 2.33), we must have $S_o = -\frac{\beta\sigma^2}{2}$, which leads to

$$A(x) = \begin{cases} 1 & \text{for } x < -\frac{\beta\sigma^2}{2} \\ e^{-\beta\left(x+\frac{\beta\sigma^2}{2}\right)} & \text{for } x > -\frac{\beta\sigma^2}{2}. \end{cases} \tag{2.47}$$

Although $\widetilde{g}(p)$ is not bounded at ∞, by a fortuitous cancellation of terms a thermal acceptance rule for the Gaussian case has been found. This cancellation was what allowed Fink & Ball [34] to spot this solution. This acceptance rule appears very similar to the Metropolis rule (eq. 1.4), and emphasises the Gaussian being a special case.

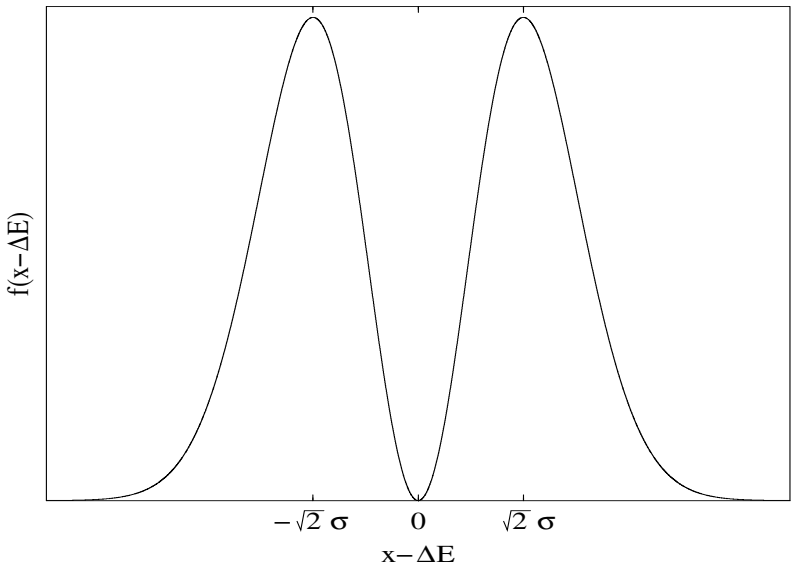

Figure 2.2: A sketch of the modified Gaussian distribution. Note that this is a bimodal distribution with Gaussian tails.

2.4.3 Modified Gaussian distribution

Now that we have a method which works well for a number of error distributions (see tables 2.1, 2.2), the question arises, does it work for all choices of f? The answer to this is a firm "no", by virtue of a distribution we have called the modified Gaussian distribution. Although this distribution will not be used for a great number of applications, it causes problems for our method. It is given by

$$f(x|\Delta E) = \frac{e^{-\frac{(x-\Delta E)^2}{2\sigma^2}}}{\sqrt{2\pi}\sigma^3}(x-\Delta E)^2 \tag{2.48}$$

where σ is a constant. This bimodal distribution is sketched in figure 2.2. The Fourier-Laplace transform of eq. 2.48 is

$$\widetilde{g}(p) = e^{\frac{\sigma^2}{2}\left(\frac{\beta}{2}+p\right)^2}\left(p+\frac{\beta}{2}+\frac{i}{\sigma}\right)\left(p+\frac{\beta}{2}-\frac{i}{\sigma}\right) \tag{2.49}$$

This leads to

$$\widetilde{\alpha}(p) = -\frac{e^{\left(\frac{\beta}{2}-p\right)S_o}}{\frac{\beta}{2}-p} + \frac{\beta}{\left(\frac{\beta}{2}\right)^2-p^2}\frac{\left(p+\frac{\beta}{2}+\frac{i}{\sigma}\right)\left(p+\frac{\beta}{2}-\frac{i}{\sigma}\right)}{1+\beta^2\sigma^2}e^{\frac{\beta\sigma^2}{2}\left(p-\frac{\beta}{2}\right)} \tag{2.50}$$

From here it is clear to see that this cannot be the transform of something which meets our restriction on $A(x)$. The problem is that $\widetilde{g}(p)$ contains a pair

Name $f(x-\Delta E)$	$A(x)$
Bimodal $\frac{1}{2}(\delta(x-\Delta E+C)+$ $\delta(x-\Delta E-C))$	1 for $x<-C$ $\frac{1+e^{\beta(C-x)}}{1+e^{2C\beta}}$ for $-C<x<C$ $\frac{2e^{\beta(C-x)}}{1+e^{2C\beta}}$ for $x>C$
Bimodal Exponentials $\frac{\gamma}{4}\left(e^{-\gamma\lvert x-\Delta E-d\rvert}+e^{-\gamma\lvert x-\Delta E+d\rvert}\right)$	1 for $x<-d$ $\frac{\gamma^2\left(1+e^{-\beta(x-d)}\right)+\beta^2 e^{-\beta(x-d)}\left(e^{-\gamma(x+d)}-1\right)}{\gamma^2\left(1+e^{2\beta d}\right)}$ for $-d<x<d$ $\frac{\gamma^2 e^{-\beta(x-d)}+\beta^2 e^{-\beta(x-d)}\left(e^{-\gamma(x+d)}+e^{-\gamma(x-d)}-2\right)}{\gamma^2\left(1+e^{2\beta d}\right)}$ for $x>d$

Table 2.2: Optimal acceptance rules for bimodal distributions - similar to the modified Gaussian distribution (eq. 2.48), but for which the method outlined in section 2.3.2 is sucessful

of zeroes for $Re(p)<0$ which is just enough to scupper our method. At this point we note that the Gaussian distribution proved troublesome, so it is perhaps not surprising that it is the modified Gaussian distribution that falls foul. Table 2.2 contains a summary of other bimodal distributions - similar to the modified Gaussian which do not suffer the same problems.

A thermal acceptance rule for the modified Gaussian distribution can be found. However, we must relinquish the insistance that $A(x)=1$ for x below some threshold. If we chose that $A(x)=1-c_1e^{-c_2x}$, where c_1 and c_2 are constants, then we find

$$A(x)=\begin{cases}1-e^{1+\left(\frac{2x}{\beta\sigma^2}\right)} & \text{for } x<-\frac{\beta\sigma^2}{2}\\ \frac{e^{-\frac{\beta^2\sigma^2}{2}-\beta x}\left(1-e^{-1-\left(\frac{2x}{\beta\sigma^2}\right)}\right)}{1+\beta^2\sigma^2} & \text{for } x>-\frac{\beta\sigma^2}{2}\end{cases}. \tag{2.51}$$

This is sketched in figure 2.3. Note that this is a rather odd acceptance rule, since it is not monotonically decreasing (as is the case for other choices of f).

Although it is satisfying that a rule can be found, this is a less than perfect situation. We have an approach, which works well in many situations. However, there is clearly a situation in which it does not work. The results reported in table 2.2 show clearly that bimodality of the distiribution is not in itself the cause of the difficulty. Currently, no way has been found to determine which distributions will work, and no more general approach which is valid for all f is known.

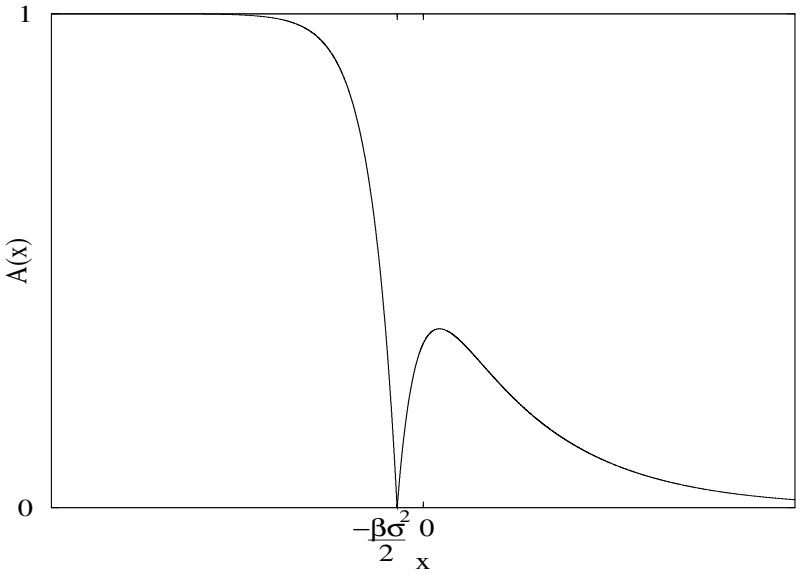

Figure 2.3: A sketch of the thermal acceptance rule for the modified Gaussian distribution. This acceptance rule is perculiar in that it is not monotonically decreasing

2.5 The width of the error distribution

Throughout, the assumption has been made that one is able to make progressively more precise estimates of ΔE. For the case of the exponential distribution, this is clearly necessary since we are not able to simulate a thermal system if $\gamma < \beta$ (see table 2.1) where γ is the width of the error distribution and β is the inverse temperature. In the Gaussian case this is not necessary, since the Gaussian distribution falls off faster than exponential. However, our method described above will become less and less efficient unless σ decreases at least as fast as β increases.

We consider the Gaussian distribution (eq. 2.2) generated from random sampling (eq. 2.7). Fink & Ball [34] addressed the issue of choosing the sample size by minimising the computational cost of accepting neutral moves (for which $\Delta E = 0$). Thus, they considered finding

$$r > 0 \text{ which minimises } \frac{r}{P_{accept}(0)}. \tag{2.52}$$

The solution of this is given by

$$r^* = 0.176\beta^2\sigma^2. \tag{2.53}$$

Here r^* is the optimal value of r for a given value of β and σ.

2.5.1 Variation of σ between moves: estimation of σ

The above would be the last word, were it not for the further complication that was mentioned earlier - that σ can vary from move to move (section 2.1). Hence, one will often have to estimate σ at the same time as ΔE. For a sample of size r drawn from a Gaussian distribution the required variance is given by $\sigma_r^2 = V_\infty / r$ where V_∞ is the variance of the population from which the estimates are drawn. For a sample of limited size, V_∞ is estimated by

$$V_r = \frac{\sum_{i=1}^{r} x_i^2}{r} - \left(\frac{\sum_{i=1}^{r} x_i}{r} \right)^2 . \tag{2.54}$$

V_r is expected to show variation around V_∞, the variance of V_r being given by

$$\sigma^2(V_r) = < V_r^2 > - < V_r >^2 \tag{2.55}$$

where $< . >$ denotes the average over all possible measurements of V_r. Writing this out in its full form and changing the order of the averaging, we find

$$\sigma^2(V_r) = \frac{\sum_{i=1}^{r}\sum_{j=1}^{r} < x_i^2 x_j^2 >}{r^2} - \frac{2\sum_{i=1}^{r}\sum_{j=1}^{r}\sum_{k=1}^{r} < x_i^2 x_j x_k >}{r^3} + \frac{\sum_{i=1}^{r}\sum_{j=1}^{r}\sum_{k=1}^{r}\sum_{l=1}^{r} < x_i x_j x_k x_l >}{r^4} - \left(\frac{\sum_{i=1}^{r} < x_i^2 >}{r} - \frac{\sum_{i=1}^{r}\sum_{j=1}^{r} < x_i x_j >}{r^2} \right)^2 . \tag{2.56}$$

For a Gaussian distribution with mean zero, this becomes

$$\begin{aligned} \sigma^2(V_r) &= V_\infty^2 \left(1 + \tfrac{2}{r}\right) - V_\infty^2 \left(\tfrac{2}{r} + \tfrac{4}{r^2}\right) + \tfrac{3V_\infty^2}{r^2} - \left(V_\infty - \tfrac{V_\infty}{r}\right)^2 \\ &= 2V_\infty^2 \left(\tfrac{1}{r} - \tfrac{1}{r^2}\right) \end{aligned} \tag{2.57}$$

So, to leading order

$$V_r = V_\infty \pm \sqrt{\frac{2}{r}} V_\infty . \tag{2.58}$$

2.5.2 How the error in σ affects the acceptance probability

This potential error in estimating σ could lead to errors in simulating a thermal system, since $A(x)$ contains a term in σ (see section 2.4.2). For small samples we expect the simulation of a thermal system to be imprecise. However, we hope that as r increases, the simulation becomes increasingly exact. For a Gaussian

error distribution (eq. 2.2) $A(x)$ is proportional to $e^{-\frac{\beta^2\sigma_r^2}{2}}$ (for $x > 0$). Since σ_r is estimated from V_r, we have

$$\frac{\beta^2\sigma_r^2}{2} = \frac{\beta^2 V_\infty}{2r} \pm \frac{\beta^2 V_\infty}{r\sqrt{2r}}. \tag{2.59}$$

We are expecting that r is chosen according to equation 2.53, which ensures that $\frac{\beta^2\sigma_r^2}{2}$ remains constant. As r increases, the portion of this term that is due to the error in σ_r decreases, and the thermal simulation becomes more precise.

It will be noted that if we estimate σ, then we may chose a non-optimal value of r. This will not affect the simulation of a thermal system, but will lead to inefficient computation.

2.6 Conclusion

We have demonstrated that it is possible to simulate a thermal system, even when the energy cannot be evaluated exactly, provided the error distribution is known. Perhaps the most interesting error distribution considered is the Gaussian case, since this corresponds to sampling errors in the central limit theorem approximation. The result for this (section 2.4.2) case is very similar to the metropolis algorithm (eq. 1.4).

This technique is slightly limited in that it gives no guarantee that $\widetilde{\alpha}(p)$ is the transform of a reasonable acceptance function. This does not pose a problem for most common error distributions. We have found one example for which we have to relinquish our insistence that $A(x) = 1$ for $x < S_o$, however, an acceptance rule for this case could still be found.

It is also interesting to look the question of how wide the error distribution should be chosen to simulate a given temperature. This is a matter of computational effieciency, rather than necessity in many cases. Fink & Ball [34] considered the Gaussian distribution and found $r^* = 0.176\beta^2\sigma^2$ to be the optimal choice of sample size. We showed that estimating σ from the same sample did not affect the thermal acceptance rule for large r.

We have considered cases when the error distribution is centred around ΔE, the exact value, but is otherwise independent of it. This method can be extended

to more general error distributions as outlined above, but this approach was not considered to be as fruitful.

Chapter 3

The Probabilistic Travelling Salesman Problem

The generation of random numbers is too important to be left to chance. – Robert R. Coveyou

Here we consider application of stochastic annealing to a particular stochastic optimisation problem, the Probabilistic Travelling Salesman Problem (PTSP). For this problem we find new results and demonstrate that stochastic annealing has significant practical utility. We begin by introducing the PTSP, and deriving some results on its asymptotic behaviour. These are compared to computational results and verified under certain limits.

3.1 Introduction

We propose the PTSP as a good test-bed amongst stochastic optimisation problems, in much the same way as the TSP has been considered a standard amongst deterministic optimisation problems. The PTSP falls into the class of NP-complete problems [16], and the TSP is a subset of the PTSP.

The traveling salesman problem (TSP) is to find the shortest tour around n cities, in which each city is visited once. For small numbers of cities this is an easy task, but the problem is NP-complete: it is believed for large n that there is no algorithm which can solve the problem in a time polynomial in n.

Consideration of the traveling salesman problem began with Beardwood et

al. [11]. They showed that in the limit of large numbers of cities which are randomly distributed on the unit square, the optimal tour length ($L_{\rm TSP}$) follows [105]

$$E(L_{\rm TSP}) = \beta_{\rm TSP}\sqrt{n} + \alpha \tag{3.1}$$

where $\beta_{\rm TSP}$ and α are constants. Here and below $E(L)$ denotes the quantity L averaged after optimisation with respect to different city positions, randomly placed on the unit square. Numerical simulation [60] gives $\beta_{\rm TSP} = 0.7211(3)$ and $\alpha = 0.604(5)$ as estimates when $n \geq 50$. Significant divergence from this behaviour is found for $n \leq 10$, but numerical estimates can be found quickly (see sec. 3.7).

The probabilistic traveling salesman problem (PTSP), introduced by Jaillet [47, 48], is an extension of the traveling salesman problem to optimisation in the face of unknown data. Whereas all of the cities in the TSP must be visited once, in the PTSP each city only needs to be visited with some probability, p. One first decides upon the order in which the cities are to be visited, the 'a priori' tour. Subsequently, it is revealed which cities need to be visited, and those which do not need to be visited are skipped to leave a 'pruned tour'. The a priori tour provides the ording in which all cities are visited, and this order must be preserved when the pruned tour is produced. The objective is to chose an a priori tour which minimises the average length of the pruned tour. Figure 3.1 shows two typical near optimal a priori PTSP tours and a typical pruned tour for each. Note that the pruned tour for $p = 0.1$ resembles an optimal TSP tour and is very different from the a priori tour.

In our terminology, the average pruned tour length is averaged over all possible instances of which cities require to be visited. This is [47]

$$\bar{L}_{\rm pruned} = p^2 \sum_{r=0}^{n-2} (1-p)^r L_t^{(r)} \tag{3.2}$$

where $L_t^{(r)}$ is the sum of the distances between each city and the following r^{th} stop on the a priori tour

$$L_t^{(r)} = \sum_{j=1}^{n} d(j, (j+r+1) \bmod n). \tag{3.3}$$

Jaillet's closed form expression for the average pruned tour length renders the PTSP to some extent accessible as a standard (but still NP complete) optimisation prob-

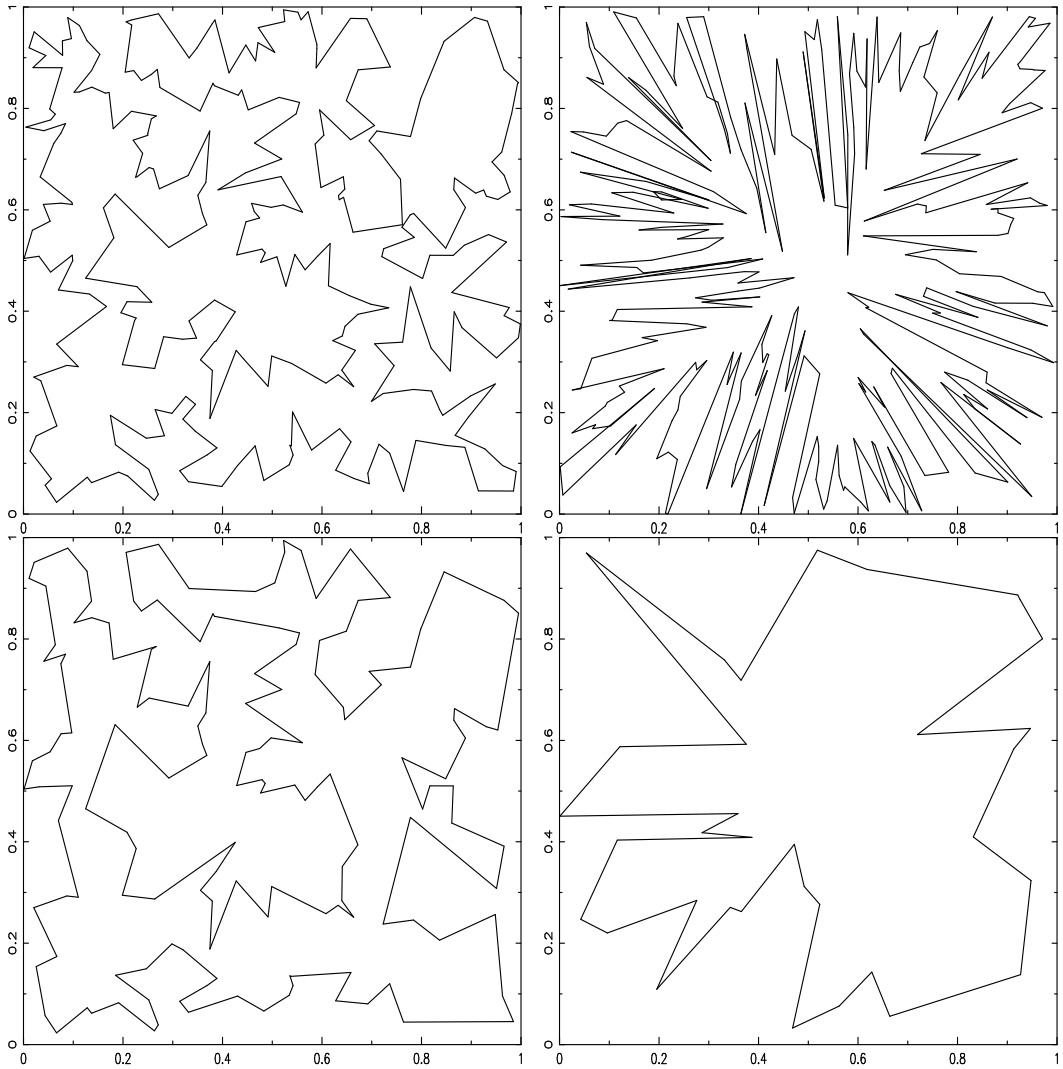

Figure 3.1: Typical near optimal a priori (top) and pruned (bottom) PTSP tours with $n = 300$ for $p = 0.5$(left) and $p = 0.1$(right), respectively. Although the a priori tour for $p = 0.1$ seems strange, the pruned tour closely resembles a TSP.

lem, and provides some check on the PTSP results by stochastic optimisation methods. For convenience $E(\bar{L}_{\text{pruned}})$, which is the expected pruned tour length further averaged over city positions after optimisation, will be referred to as the expected pruned tour length.

It has been conjectured [16] that, in the limit of large n, the PTSP strategy is as good as constructing a TSP tour on the cities requiring a visit, the re-optimisation strategy. The re-optimisation strategy gives

$$\lim_{n\to\infty}\left(\frac{E\left(\bar{L}_{\text{Reopt}}\right)}{\sqrt{np}}\right) = \beta_{\text{TSP}}. \tag{3.4}$$

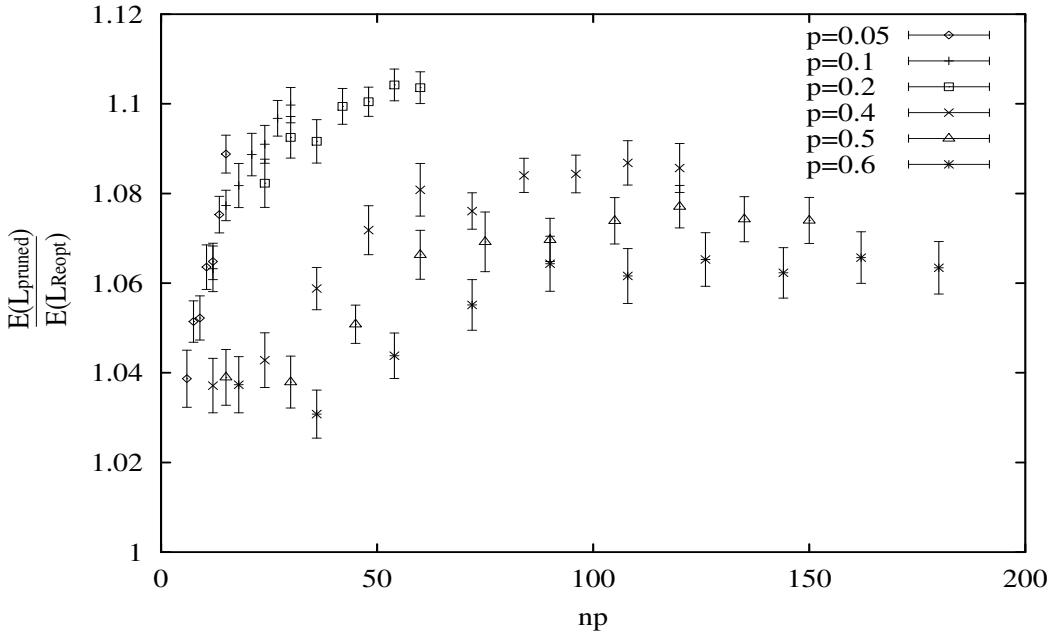

Figure 3.2: The expected pruned tour length divided by the expected re-optimised tour length, indicating the improvement one would expect from re-optimisation. These results were obtained using stochastic annealing as outlined in sec. 3.3. The expected re-optimised tour length is calculated using eq. 3.2 with $L_t^{(r)}$ replaced by the expected optimal length of a TSP tour over r cities.

Figure 3.2 shows the expected pruned tour length divided by the expected re-optimised tour length. Since this quantity is tending towards a value significantly greater than 1 for $p < 1$ it demonstrates that the PTSP strategy can be worse than the re-optimisation strategy. Jaillet [47] and Bertsimas et al. [15] have shown that

$$\lim_{n\to\infty}\left(\frac{E\left(\bar{L}_{\text{pruned}}\right)}{\sqrt{np}}\right) = \beta_{\text{pruned}}(p) \tag{3.5}$$

where

$$\beta_{\text{TSP}} \leq \beta_{\text{pruned}}(p) \leq Min(0.9212, \frac{\beta_{\text{TSP}}}{\sqrt{p}}). \tag{3.6}$$

One attempt to solve the PTSP using an exact method was taken by Laporte et al. [58] who introduced the use of integer linear stochastic programming. Although use of algorithms which may exactly solve the PTSP are useful, they are always very limited in the size of problem which may be attempted. Furthermore, the stochastic programming algorithm even fails to solve the PTSP on certain occasions, thus the accuracy of any statistics that would be generated using this method is dubious.

Three studies have used heuristics to solve the PTSP [13, 94, 14]. None of these studies used global search heuristics, and were very restricted in the problem size attempted because equation 3.2 takes $O(n^2)$ time to calculate. Thus, to solve a 100 city problem for the PTSP would take $O(10,000)$ times longer than it would to solve a 100 city problem for the TSP. It should be noted, however, that it is only possible to make this comparison due to the relative simplicity of the PTSP. For many more stochastic optimisation problems, standard optimisation techniques are not applicable.

3.2 Form Of The Optimal Tour & Scaling Arguments

Optimal a priori PTSP tours for small p, as exemplified in figure 3.1 for $p = 0.1$, resemble an "angular sort" - where cities are ordered by their angle with respect to the centre of the square. Bertsimas [13] proposed that an angular sort is optimal as $p \to 0$, but we can show this to be false by comparison to a space-filling curve algorithm which is generally superior as $n \to \infty$. Such an algorithm was introduced by Bartholdi et al. [8] using a technique based on a Sierpinski curve.

For the angular sort with $np \gg 1$, the probability of two cities being nearest neighbours on the pruned tour will be vanishingly small for cities which are separated from each other by a large angle on the a priori tour. This means that only cities that are separated by a small angle contribute significantly to eq. 3.2. Thus for an n city tour, we may approximate

$$L_t^{(r)} \simeq L_o n \tag{3.7}$$

where L_o is some fraction of the side of the unit square, since cities which are sorted with respect to angle will be unsorted with respect to radial distance. This leads to

$$E(\bar{L}_{\mathrm{ang}}) \simeq L_o n p^2 \sum_{r=0}^{n-2} (1-p)^r. \tag{3.8}$$

For $np \gg 1$, we find

$$E(\bar{L}_{\mathrm{ang}}) \to L_o n p. \tag{3.9}$$

By contrast it has been shown [15] that

$$\frac{E(\bar{L}_{\tau_{\mathrm{sf}}})}{E(\bar{L}_{\mathrm{Reopt}})} = C \tag{3.10}$$

with probability 1, where $E(\bar{L}_{\tau_{\rm sf}})$ is the expected length of a pruned tour generated by a heuristic based on the Sierpinski curve and $E(\bar{L}_{\rm Reopt})$ is the expected pruned tour length for the re-optimisation strategy. Using previous computational results [15, 60], we estimate $C \simeq 1.33$, which is worse than we achieve using stochastic annealing. Hence, $E(\bar{L}_{\tau_{\rm sf}})$ is given by

$$E(\bar{L}_{\tau_{\rm sf}}) = O(\sqrt{np}) \tag{3.11}$$

which leads to

$$\frac{E(\bar{L}_{\tau_{\rm sf}})}{E(\bar{L}_{\rm ang})} = O(\frac{1}{\sqrt{np}}). \tag{3.12}$$

So, for any $p > 0$, provided n is sufficiently large, the angular sort is not optimal.

From diagrams of near-optimal PTSP tours such as fig. 3.1, we propose that the tour behaves differently on different length scales; the tour being TSP-like on the larger length scales, but resembling a sort at smaller distances. We may construct such a tour and use scaling arguments to analyse the expected lengths of both the pruned and a priori tours for optimal solutions of the PTSP.

Consider dividing the unit square into a series of 'blobs', each blob containing $1/p$ cities so that of order one city requires a visit. The number of such blobs is given by

$$N \simeq np \tag{3.13}$$

and for these to approximately cover the unit square their typical linear dimension ξ must obey

$$N\xi^2 \sim 1. \tag{3.14}$$

Since each blob is visited of order once by a pruned tour, we can estimate the expected pruned tour length to be

$$E(\bar{L}_{\rm pruned}) \sim N\xi \sim \sqrt{np} \tag{3.15}$$

which we will see below is verified numerically. We can similarly estimate the a priori tour length to be n times the distance between two cities in the same blob. Thus, the expected a priori tour length is

$$E\left(L_{\rm a\ priori}\right) \sim n\xi \sim \sqrt{\frac{n}{p}} \tag{3.16}$$

which is more difficult to confirm numerically.

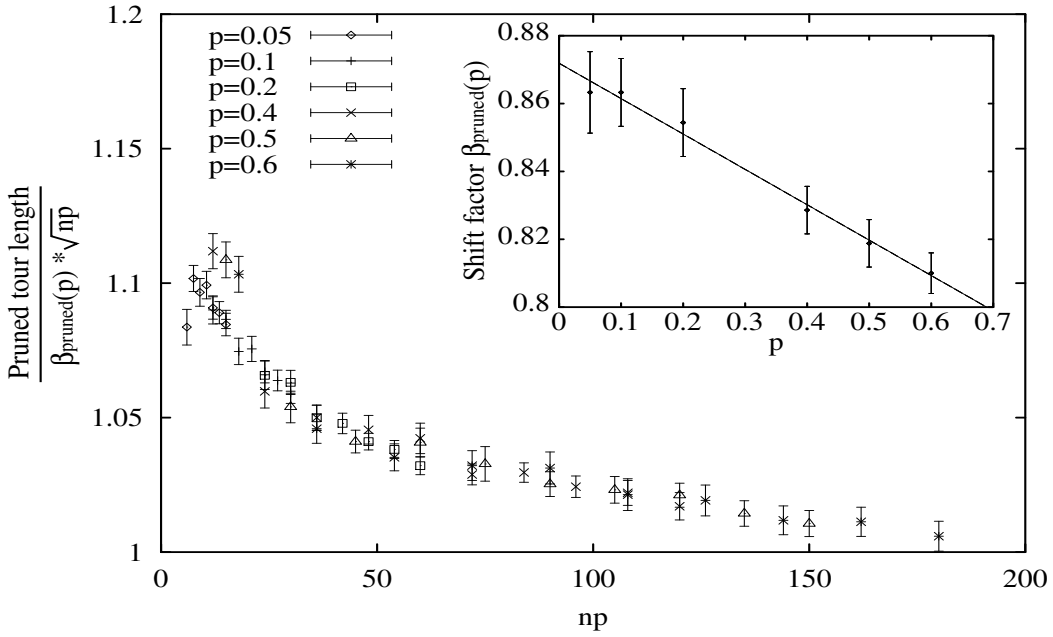

Figure 3.3: The master curve for the pruned tour length divided by $\beta_{\text{pruned}}(p)\sqrt{np}$. The data follows a smooth curve for $np > 30$, and the shift factors follow a linear relationship, suggesting that $\frac{E\left(\bar{L}_{\text{pruned}}\right)}{\sqrt{np}(0.872-0.105p)} = f(np)$.

3.3 Computational Results For The PTSP

We have investigated near optimal PTSP tours for a range of different numbers of cities, and various values of p. Effective temperatures in the range $kT = 0.07 - 0.01$ were used, giving sample sizes in the range $r = 2 - 500$. Between 10 and 80 different random city configurations were optimised (80 configurations of 30 cities, 40 configurations of 60 cities, 20 configurations of 90 cities and 10 configurations for $n \geq 120$ cities).

Figure 3.3 shows a master curve for the expected pruned tour length divided by $\sqrt{np}$. The shift factors have a linear fit and the data are consistent with

$$\frac{E(\bar{L}_{\text{pruned}})}{\sqrt{np}(a - bp)} = f(np) \tag{3.17}$$

for $n \gg 1$, where $a = 0.872 \pm 0.002$, $b = 0.105 \pm 0.005$ and $f(np) \to 1$ for large np. The shift factors indicate that the PTSP strategy can be no more than $\frac{0.872}{0.767} - 1 = 14\%(\pm 1\%)$ worse than the re-optimisation strategy.

The master curve for the a priori tour length is shown in fig. 3.4. Our scaling arguments predict that the shift factors $\beta_{\text{a priori}}(p)$ should tend towards a constant

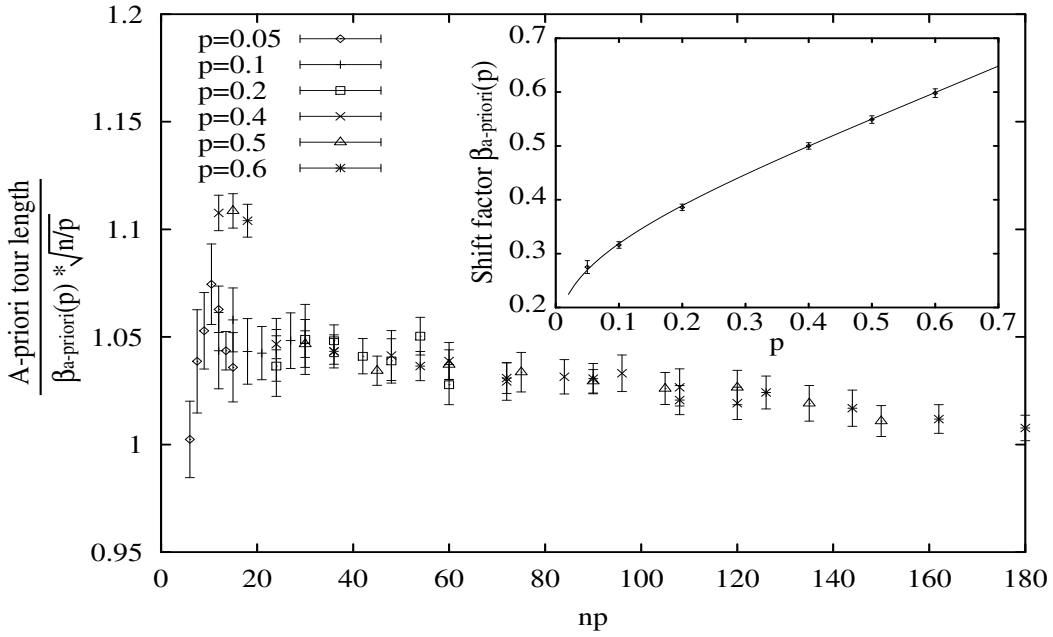

Figure 3.4: The master curve for the a priori tour length divided by $\sqrt{\frac{n}{p}}\beta_{\text{a priori}}(p)$. The shift factors, inset, are expected to tend towards a constant for $p \to 0$. The slight, but significant, deviation from linear suggests that this might not be the case.

for $p \to 0$. However, data are fit very well by the relation

$$\beta_{\text{a priori}}(p) = \frac{1}{1.25 - 0.82ln(p)} \tag{3.18}$$

which would tend to zero as $p \to 0$ in conflict with our scaling arguments. To resolve this dilemma we need to probe very small p.

3.4 The Limiting Case $p \to 0$

We are interested in finding whether $\beta_{\text{a priori}}(p)$ tends towards a constant as $p \to 0$. To do this using the above approach is difficult, since we need a large number of cities to produce reliable data for this regime. Extraction of this behaviour can be achieved by comparing simulations for different values of n, but fixed np. We accomplish this by insisting that each instance has 4 cities on the pruned tour. 4 city tours are chosen since they are the smallest for which it matters in which order the cities are visited. This can be viewed as an efficient way to simulate (approximately) the PTSP strategy with $p = \frac{4}{n}$.

Since we are considering PTSP's at fixed np, if $\beta_{\text{a priori}}(p)$ tends towards

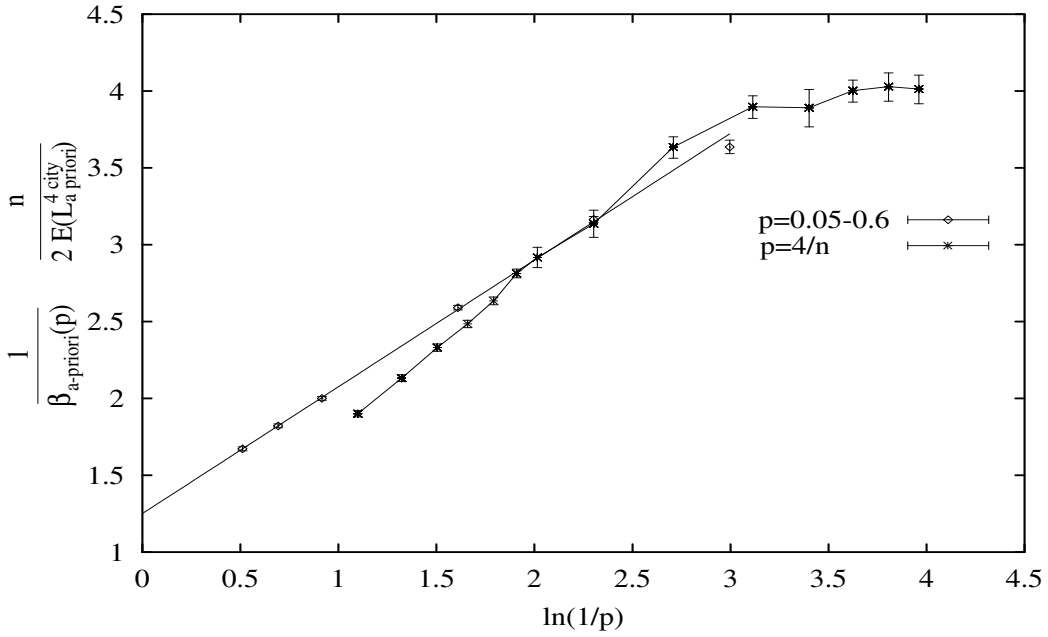

Figure 3.5: Shift factors for a priori tours compared to the equivalent measurement of 4 city tours. The 4 city tour data is optimised when each of the instances have 4 cities on the pruned tour. The diamond points show $\frac{1}{\beta_{\text{a priori}}(p)}$ does not appear to saturate within the accessible range of p. The stars show matching behaviour, with saturation at larger n corresponding to inaccessible p, suggesting that $E(L_{\text{a priori}}) = \beta\sqrt{\frac{n}{p}}$ for small p.

a (non-zero) constant as $p \to 0$ then by eq. 3.16 we expect $E\left(L^{4\text{ city}}_{\text{a priori}}\right)/n$ to tend towards a constant as $p \to 0$. Simulations in this regime were performed for $n = 12 - 210$. 100 different random city configurations were used for $n < 30$, 20 configurations were used for $n \leq 90$ and 10 configurations for $n \geq 120$. Figure 3.5 shows a linear-log plot of $\frac{n}{2E\left(L^{4\text{ city}}_{\text{a priori}}\right)}$ against $\ln(n/4) \cong \ln(1/p)$. For small n these results reasonably match the direct measurements of $\beta_{\text{a priori}}(p)$ shown for comparison. For surprisingly large $n \sim 100$, beyond the range of our $\beta_{\text{a priori}}(p)$ data, $E\left(L^{4\text{ city}}_{\text{a priori}}\right)/n$ does approach a constant value, consistent with our earlier proposal of scaling behaviour. To summarise

$$E(L_{\text{a priori}}) = \sqrt{\frac{n}{p}}\beta_{\text{a priori}}(p) \tag{3.19}$$

where

$$\beta_{\text{a priori}}(p) \begin{cases} = \frac{1}{1.25 - 0.82\,ln(p)} & p > 0.03 \\ = \beta_0 & p < 0.03. \end{cases} \tag{3.20}$$

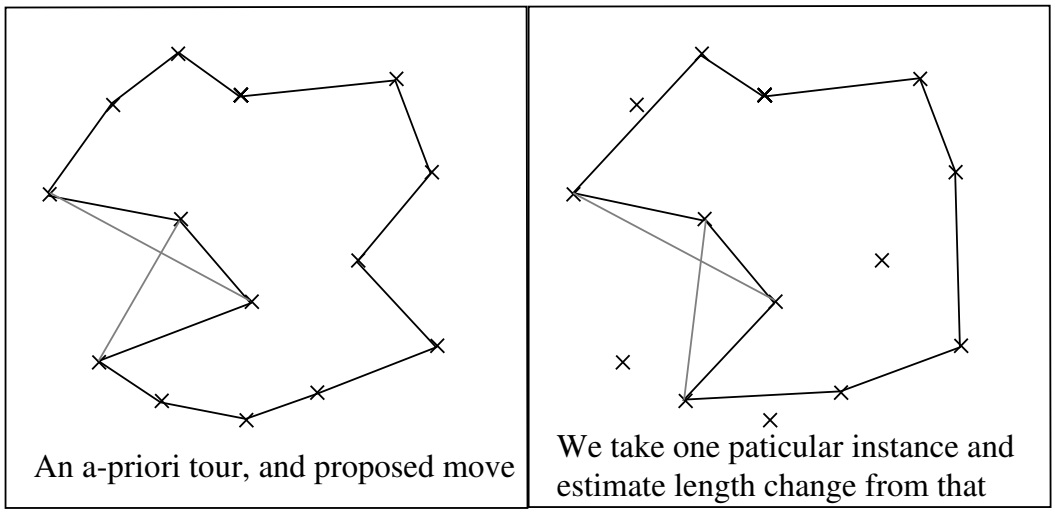

Figure 3.6: When estimating the expected length change due to a move, we randomly generate instances of which cities require a visit. Only the cities that are nearest to the move are needed to calculate the calculate the change in the pruned tour length.

3.5 Notes on Algorithm Implementation

Stochastic annealing was implemented with a combination of the 2-opt and 1-shift move-sets [63]. Both move-sets work similarly to that which would be expected for the deterministic case. The expected pruned tour length change for the move is estimated by averaging the change in the tour length for a number of instances. For a given instance of which cities require a visit it is not necessary to decide whether every city is present, but only the set of cities closest to the move - that determine the change in the tour length (see figure 3.6). For the PTSP, the location of the nearest cities on the pruned tour to the move is determined from a simple Poisson distribution.

When using stochastic optimisation, the only variable over which we have control is the sample size (the number of instances), whereas the temperature is $\frac{\sigma}{\sqrt{r}}$. As shown in figure 3.7, annealing by controlling r alone exhibits a relatively sharp transition in the expected pruned tour length. The rapid transition appears to 'freeze in' limitations in the tours found (analogous to defects in a physical low temperature phase). Compare this to the smooth change seen when $\frac{\sigma}{\sqrt{r}}$ is controlled. The system was annealed at each value of the temperature and value of r for 50,000 Monte Carlo steps with $n = 300$ and $p = 0.1$.

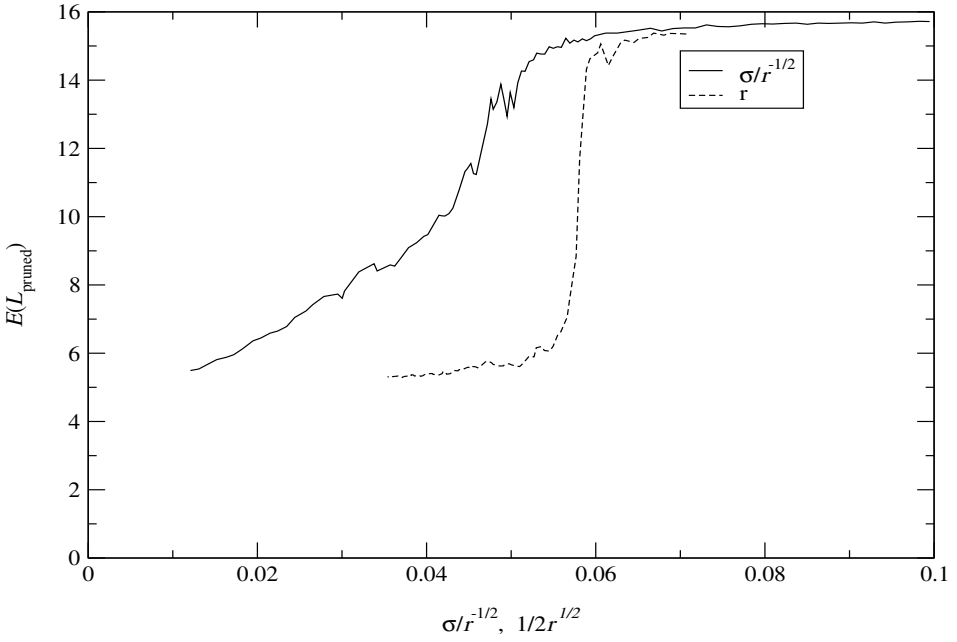

Figure 3.7: The expected pruned tour length for annealings when r and $\frac{\sqrt{r}}{\sigma}$ are increased monotonically. The jump in the pruned tour length is seen when only r is controlled, demonstrating that this "freezes in" imperfections in the tour.

The transition is caused by the fact that σ may vary from move to move, and is on average lower when the expected pruned tour length is less. The jump in the pruned tour length is accompanied by a jump in σ and hence the temperature. We suggest that quite generally controlling $\frac{\sigma}{r}$ gives a better cooling schedule than focussing on r alone.

3.6 Conclusion

We have considered the PTSP as a difficult stochastic optimisation problem and introduced a crossover scaling interpretation of the PTSP where the a priori tour is TSP-like on large scales, and a sort on shorter length scales. This gives behaviour which agrees with the computational results, with the pruned tour length given by eq. 3.17.

The a priori tour length is more subtle in nature than the pruned tour length. We introduced 4-city tours to probe the a priori tour length as $p \to 0$. This allows us to conclude that the crossover scaling interpretation gives the correct behaviour

for $p < 0.03$, as given by eq. 3.19. Direct confirmation of this result is left as a future challenge.

3.7 Appendix

Numerical estimates of the length of a TSP tour for $n \leq 10$ are given below

Number of cities n	Number of different city configurations (I)	Average tour length	$\sigma/\sqrt{I-1}$
2	100000	1.043429	0.002
3	100000	1.564702	0.002
4	5000	1.889601	0.006
5	5000	2.123484	0.006
6	5000	2.311458	0.005
7	5000	2.472799	0.005
8	5000	2.616990	0.005
9	5000	2.740075	0.005
10	5000	2.862946	0.005

Table 3.1: The expected length of the optimal TSP tour for small numbers of cities, averaging over I different random city configurations.

Chapter 4

Oil Field Optimisation

Where there's muck there's brass. – Anon.

To demonstrate the utility of stochastic annealing (and stochastic optimisation) we consider its application to oil field projects. These projects are characterised by a large uncertainty in many of the parameters. In this example, we consider the primary source of uncertainty to be the reservoir size.

4.1 Introduction

Oil fields are the subject of intense study. Each development can be worth many billions of dollars in revenue. Hence, any small improvement in strategy is worth vast sums of money.

Given this, one might expect that every element associated with oil field development was a precise art. However, estimating the size of a reservoir remains very imprecise. Since the uncertainty is so large, this provides a good demonstration of the applicability of stochastic annealing to a commercial problem.

4.1.1 The creation of oil

Petroleum (oil is liquid petroleum) is formed from the decay of organic matter as it is exposed to high temperatures and pressures underground ([72], vol. 13, PP. 290-291). Specific conditions are required so that the organic matter does not rot, burn or decompose into other chemicals.

Extraction of oil from the rocks in which it was created is often difficult and expensive, so most of the oil currently recovered has migrated from the source rocks to a suitable reservoir. A good reservoir is a volume of porous rock surrounded by non-permeable rock in such a way that it is difficult for the oil to escape.

The rocks in which oil resides must also be permeable, so that the oil can be extracted at a tolerable rate. If the oil is highly pressurised, it may be extracted under its own pressure. However, in most cases the oil must be lifted to the surface by some means. Often these means are less than efficient, and it is estimated that only 35% of the total oil currently found in the world will be extracted ([72], vol. 13 P. 291).

4.1.2 Oil field discovery

Once an area has been found with the appropriate geological conditions for the formation of oil, more detailed surveys can be performed to search for the presence of oil.

Detailed information can be provided by seismic surveying ([72], vol. 16, PP.218-221). A disturbance is created (by an explosion, for example) at some point, and movements of the earth are measured at an array of points (see fig. 4.1). The waves travel at different speeds through different parts of the earth, and are partially reflected at boundaries between volumes which have different acoustic properties. The detectors are able to measure the reflected waves and (in principle) extract the properties of the sub-surface. However, seismic measurements are subject to a great deal of noise, and it may not be possible to uniquely identify the cause of a given feature in the data. In particular, it is difficult to distinguish between a reservoir containing oil and one containing water.

The final tool for the discovery of oil is the drilling of wells. Even if seismological surveys identify a reservoir where oil may be present, there is still no guarantee that oil is to be found in this place, or that it is possible to extract this oil. The only definitive method for establishing the presence of oil at a given point is the drilling of a well. A series of wells are required to delimit the size of a given reservoir and to test whether the oil is split among a number of compartments. Since the drilling of wells is very expensive, as few as possible are drilled for this purpose. This leaves

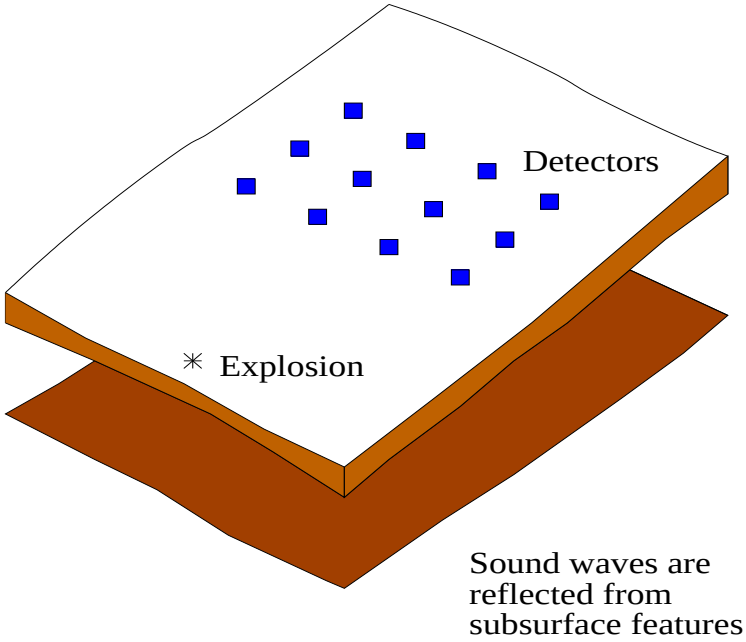

Figure 4.1: The typical setup for a (3D) seismic survey. An explosion creates a pressure wave which propagates through the earth. These waves are affected by changes in the properties of the earth, and reflections from sub-surface features will be measured by an array of detectors.

a large uncertainty in the size of a given reservoir, and means that oil production is an inherently risky business.

4.2 Uncertainty

When production decisions are made, there can still be a large amount of uncertainty in some of the key parameters of the oil field.

4.2.1 Production uncertainties

Each oil field will normally contain a number of reservoirs. Some of these will have been thoroughly investigated, but with significant uncertainty in their size still remaining. Some will have been seismically surveyed, but not drilled, and therefore may not contain oil.

We consider the situation where the decision to produce oil has been taken, but where much uncertainty remains. The following variables are assumed to be

uncertain:

- Reservoir size (including whether oil is present)
- The cost of drilling a production well
- Facilities uncertainties :
  - Sub-sea cost
  - The cost per km of a pipeline
  - The cost of vessels and installations
  - The cost of risers
- Operating expenditure
- Capital operating expenditure
- Ratio of the cost of drilling an appraisal well to a production well

4.2.2 Oil price fluctuations

In addition to uncertainties in geological and production parameters, variability in financial factors also plays an important role [78]. It has been claimed that fluctuations in the oil price are the most significant uncertainty [66]. Figure 4.2 shows the price of crude oil for the years 1974-2000. The fluctuations are severe, showing a few large swings in the price. This uncertainty has been considered by Jonsbraten [50] where a small number of differing scenarios for the future oil price were taken. The effect of variability in interest rates (the discount factor) has also been considered [66], but was found to be less important.

As was pointed out by Paddock et al. [88], oil field projects are ones for which a number of decisions are taken at different times and this is an important consideration when valuing oil fields. An oil field may only be profitable when the oil price is high, and hence exploration may only be a viable commercial decision under those conditions [30]. Further, only producing oil at times when prices are high has been considered [61].

In this study (for simplicity) we will follow the widespread practice in the oil industry of optimising under a fixed oil price and discount factor, focussing on the impact of geological factors. Although choosing a fixed, low oil price ensures that the company is not susceptible to many fluctuations in oil price and allows for clear comparison between projects, the strategy for each project may not be optimal.

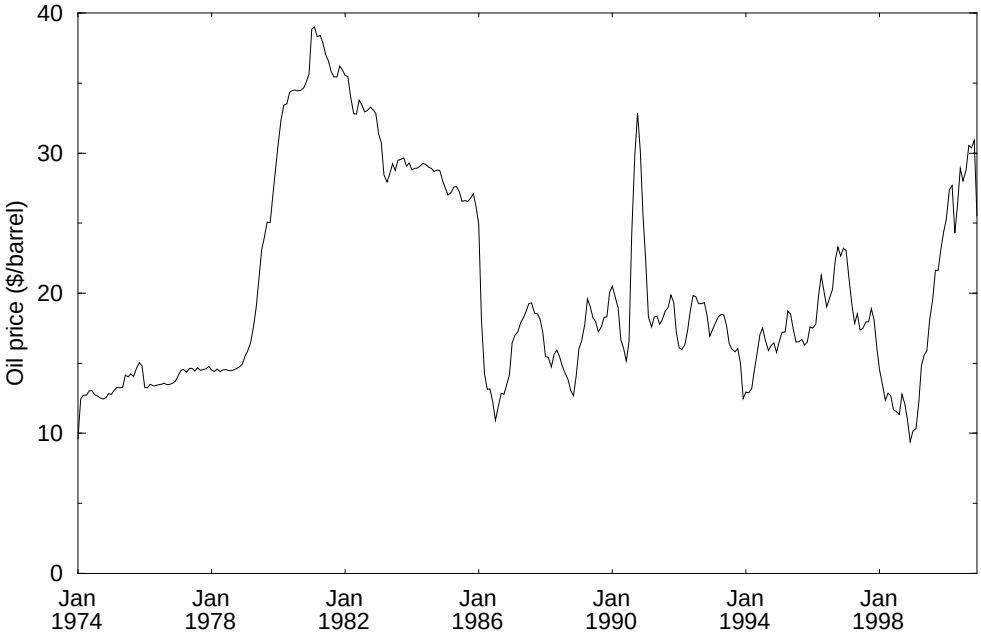

Figure 4.2: The price of crude oil (in $/barrel) from 1974 to 2000. Since we do not know currently know the price of oil in the future, models for oil field development should account for the fluctuation in the future oil price.

4.2.3 Reducing risk

Given the large uncertainty in the size of the reservoirs, one may wish to conduct extra exploration to determine their size more accurately. With the stochastic annealing approach, we are able to evaluate the extra worth of a given exploration. For example, consider a survey on a reservoir, which is able to pinpoint the exact size of that reservoir. We consider such an investigation on reservoir 1. We may write the standard stochastic optimisation problem as

$$\text{find } \mathbf{x_o} \text{ which maximises } \bar{g}(\mathbf{x_o}) = \int\int g(\mathbf{x_o}, \omega, \omega_1) P(\omega) P_1(\omega_1) d\omega d\omega_1 \qquad (4.1)$$

where $P_1(\omega_1)$ is the probability of reservoir 1 having size ω_1, ω represents all the other unknowns in the problem and $g(\mathbf{x_o}, \omega, \omega_1)$ is the value of strategy $\mathbf{x_o}$. If we determine the value of ω_1 before optimisation, then we will be seeking to optimise the following

$$\text{find } \mathbf{x_1}(\omega_1) \text{ which maximises } \bar{g_1}(\mathbf{x_o}, \omega_1) = \int g(\mathbf{x_o}, \omega, \omega_1) P(\omega) d\omega. \qquad (4.2)$$

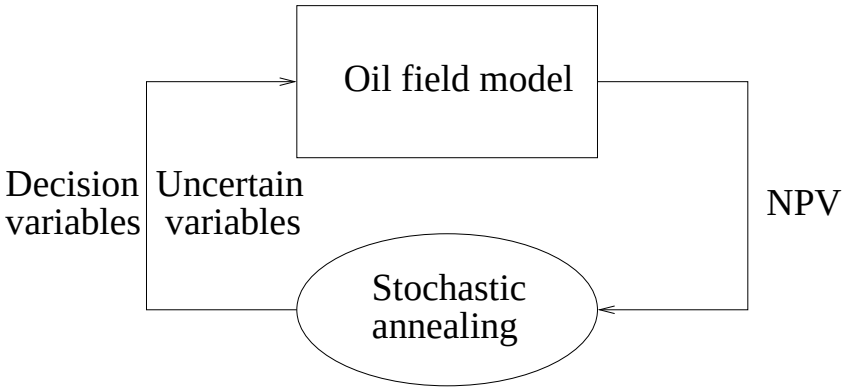

Figure 4.3: The process of optimisation. The oil field model requires a set of decision variables, and values for the variables which are uncertain. Given these it produces a value for the NPV. The details of the model are not important, and stochastic annealing can optimise any oil field model.

The expected value of the entire project will thus be given by

$$EV = \int \bar{g}_1(\mathbf{x_1}, \omega_1) P_1(\omega_1) d\omega_1 \tag{4.3}$$

Providing the information is accurate, then $EV \geq \bar{g}(\mathbf{x_o})$. However, the exact value of a piece of information is an important quantity.

4.3 The oil field model

The key elements of the oil field model are outlined below. Since the model has been developed for commercial use by BP [102] it is neither possible nor desirable to fully describe it. The model is highly complex and tries to be an accurate representation of the development. As far as possible the model has been treated as a "black box", the details of which are irrelevant for the optimisation and could, in principle, be replaced by any other oil field model. Figure 4.3 shows a flow diagram for the process of optimisation.

- The oil is lifted from the ground by water injection. Water is pumped into the reservoir to maintain a constant pressure underground. A typical production curve is shown in figure 4.4.

- The number of each type of well to be drilled into a given reservoir is fixed a-priori, with no possibility for multi-stage drilling.

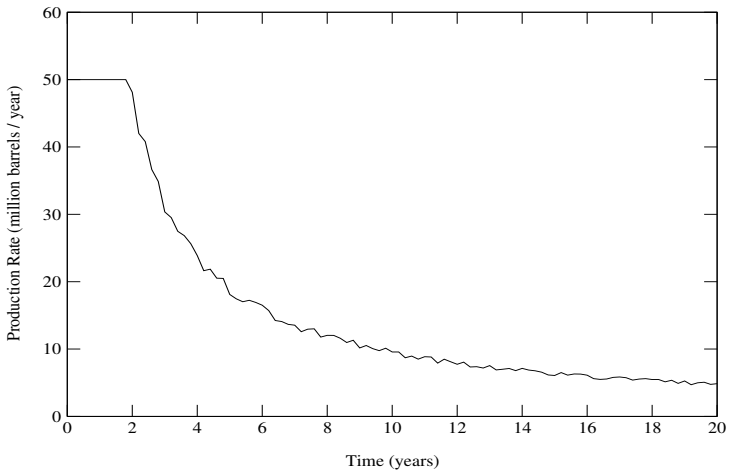

Figure 4.4: A typical production rate for a reservoir. Oil is produced at a constant rate (limited by facilities) for a given time, after which production drops off sharply due to water breakthrough, or pressure decline.

∘ Reservoirs naturally appear together in clusters. All the wells that are to be drilled in a given cluster must be drilled at the same time, for convenience of drilling facilities.

∘ Oil is pumped from each cluster to a hub, and each cluster is "tied in" individually.

∘ The hub location is restricted, since it is expensive to pump oil up a steep hill. The hub is restricted to a given line in the model.

∘ Oil is sold at the fixed price of \$15 per barrel.

∘ The distribution of the size of a reservoir is shown in figure 4.5.

∘ The distribution of the other uncertain variables is taken to follow a triangular distribution of width $\pm$ 20% around the modal value.

∘ The variables which are optimised are: the order in which the clusters are drilled, which reservoirs within each cluster should be drilled, the hub location and the hub capacity.

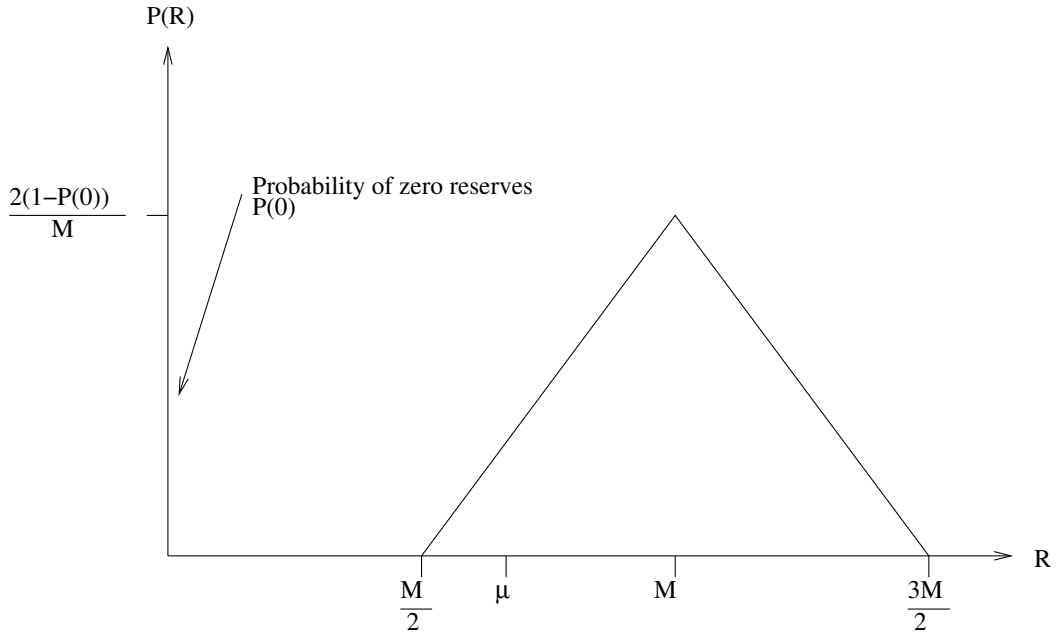

Figure 4.5: The distribution of the size of a reservoir $P(R)$. There is a finite probability that there are no reserves $P(0)$. If there is oil present, the size of the reservoir is given by a triangular distribution around the modal size of the reservoir, M. The mean value of the reserves μ is less than the modal value, and given by $\mu = (1 - P(0))M$.

4.3.1 Optimisation method

This study has been made possible by the availability of a commercial model of an oil field which exists in a form amenable to optimisation methods. This model is composed of 26 linked worksheets. Due to the size of the model it was not feasible to export it into computer code, which would bring a significant improvement in speed. Hence, only limited optimisations could be performed.

The optimisation technique was based on the Fink & Ball [34] algorithm where all apparently favourable moves are accepted. Thermal acceptance rules (as detailed in sec. 2.4) were not used, as they lead to a greater number of moves being rejected. To explore temperatures higher than given by the raw Fink & Ball algorithm (eq. 2.6) for a sample size of 1, a hybrid algorithm was devised. For the two different optimisations (sec. 4.4.2 and 4.4.6) two different hybrid algorithms were used.

In the first case (sec. 4.4.2), moves that appeared to be unfavourable were accepted with a probability of $e^{-\beta x}$ where x is the apparent energy change. β was

chosen as $\frac{1}{\beta} = 10,000$ to simulate high temperatures, and $\beta = 1$ otherwise. The sample size was chosen to be $r = 1, 10, 100, 1000$ as the annealing progressed.

In the second case (sec. 4.4.6), all moves for which the apparent energy change was less than some threshold were accepted. This threshold was chosen sequentially to be 10,000, 1,000, 100 and 0. The sample size was then chosen to increase from 1 to 10 and then 100. The best policies at this stage were then quenched using a sample size of 1,000.

The reported results are expected values of the financial measures estimated from a sample of size 1,000.

4.4 Results

4.4.1 What is the appropriate figure to optimise?

Once we have a model of the production of an oil field, we must choose an output quantity which we wish to optimise. The most obvious is the net present value (NPV). This is

$$NPV(d) = \sum_{i=0}^{\infty} \frac{V(i)}{(1+d)^i} \tag{4.4}$$

where the oil field makes an average of $V(i)$ million dollars profit (net cash inflow) i years from now and d is the discounting rate. A common figure to optimise is $NPV(0.1)$, since it is believed that the value of stocks increases by 10% per year on average.

Since investors want to protect themselves from risk, it is important to consider the variability in the NPV. We may therefore wish to ensure that $\sigma(NPV(d))$ is not too large, which is a simple quantity to estimate by Monte Carlo sampling. Hence we may propose to optimise the function

$$U = NPV(d) - k\sigma(NPV(d)) \tag{4.5}$$

where k is some constant. In the case of our oil field model $d = 0.1$ and $k = 0$ is chosen, no account being taken of uncertainty.

All the financial quantities reported here have been multiplied by an arbitrary factor to maintain confidentiality of information. This, however, does not alter any comparisons.

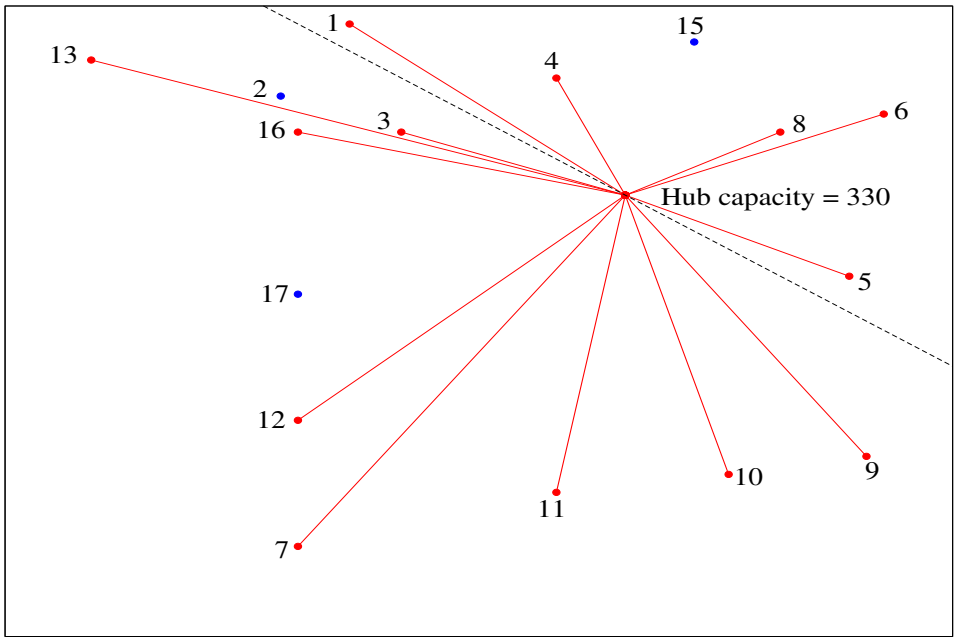

Figure 4.6: The configuration of the optimal policy as found by stochastic annealing. Each cluster is given a number, and those which are shown as connected to the hub are those which will be drilled. The hub is restricted to be on the line shown. The expected NPV of this policy is $NPV(0.1) = 1455$.

4.4.2 Optimal policy

The best policy found by stochastic annealing is illustrated in figure 4.6. This policy develops almost all the clusters, which can be thought of as meaning that the rewards for a reservoir being large are greater than the penalty for it being small (or empty). Also, since the quantity being optimised is risk neutral ($k = 0$), this policy may contain a number of clusters which are only marginally profitable, but increase the risk noticeably.

This policy is found to have the following financial measures

$$\begin{aligned} NPV(0) &= 8470 & \sigma(NPV(0)) &= 2029 \\ NPV(0.1) &= 1455 & \sigma(NPV(0.1)) &= 585 \\ IRR &= 25.0\% & \sigma(IRR) &= 3.6\% \end{aligned} \tag{4.6}$$

where IRR is the internal rate of return, given by $NPV(IRR) = 0$.

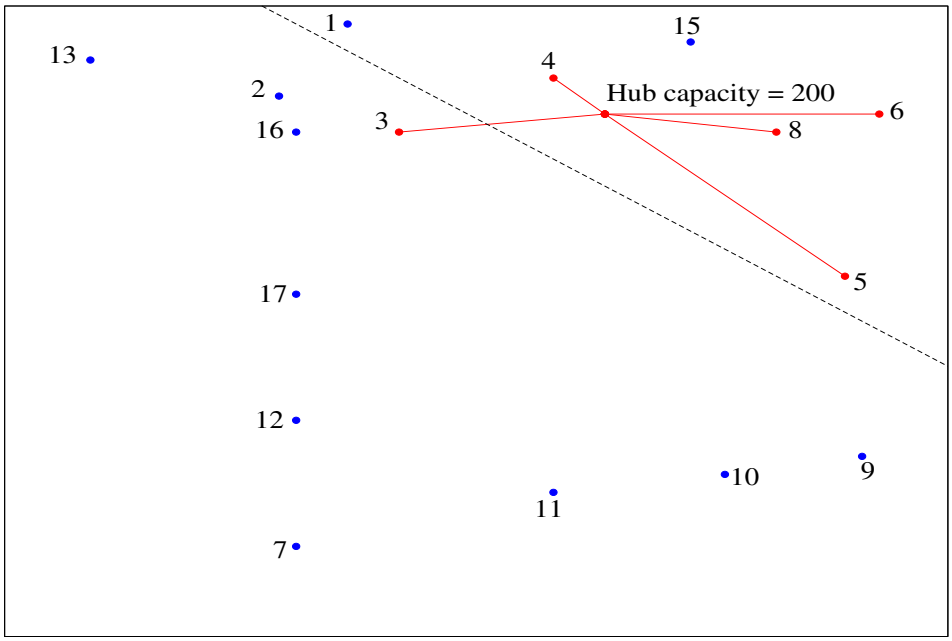

Figure 4.7: The configuration of the base case (which was proposed by BP). This is a very conservative policy, with 5 clusters being drilled. The hub is only restricted to the line shown for the optimisation, as this aids computational simplicity. The expected NPV of this policy is $NPV(0.1) = 952$.

4.4.3 The base case

We compare the strategy given by stochastic optimisation with the "base case". This strategy is one developed by BP. Hence, it will take the uncertainty into account (at least in part).

This policy is much more conservative and involves developing only 5 clusters, see figure 4.7. It has the following financial measures

$$\begin{aligned} NPV(0) &= 5245 & \sigma(NPV(0)) &= 1354 \\ NPV(0.1) &= 952 & \sigma(NPV(0.1)) &= 408 \\ IRR &= 25.3\% & \sigma(IRR) &= 4.0\% \end{aligned} \tag{4.7}$$

The optimal policy achieves a value of $NPV(0.1)$ 50% greater than the base case. This shows that the base case is cautious - to the point of being a bad policy. If we consider the measure $\frac{\sigma(NPV(0.1))}{NPV(0.1)}$ as an expression of the relative risk of a project, we find that the optimal policy has a lower value of this measure.

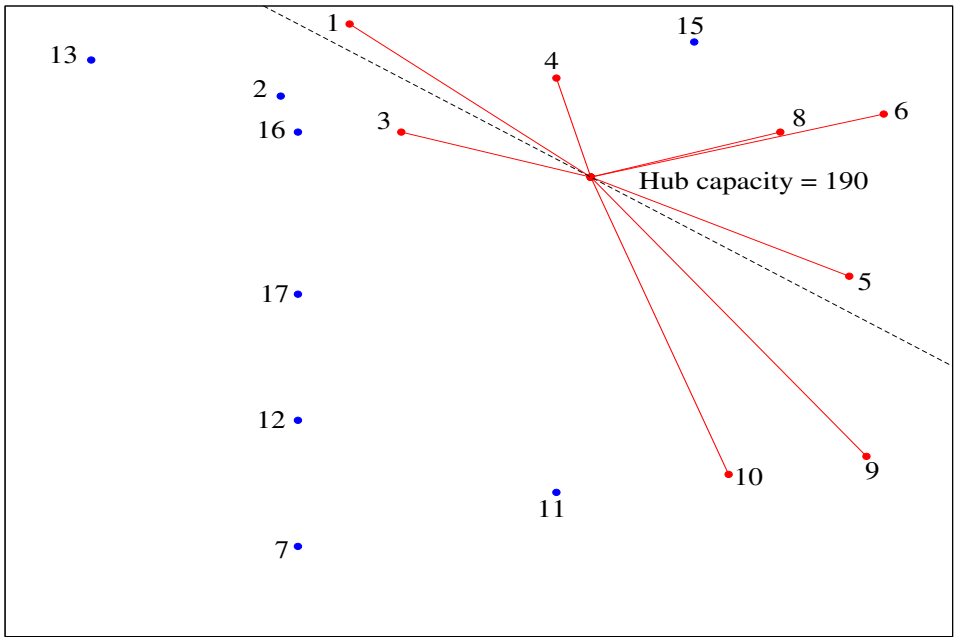

Figure 4.8: The configuration of the policy chosen by optimisation when the size of each reservoir is assumed to be known. This policy develops 8 clusters and is more optimistic than the base case. The expected NPV of this policy is $NPV(0.1) = 966$.

4.4.4 Deterministic optimisation

We now make a comparison between the strategy given by stochastic optimisation, and one where the size of each reservoir is assumed to be known. We choose the size of each reservoir to be given by the average of the distribution of sizes assumed by the stochastic version of the model. This is expected to perform well, but the extra assumption of deterministic reserves may affect its strategy. The solution was found by performing a number of optimisations using simulated annealing with a Metropolis rule (see eq. 1.4), the temperature ranging from $\frac{1}{\beta} = 10,000$ to $\frac{1}{\beta} = 1$ in steps of factor 10.

The policy given by optimisation under deterministic reserves gives a more optimistic policy than the base case and is illustrated in figure 4.8. The financial measures are

$$\begin{aligned} NPV(0) &= 6102 & \sigma(NPV(0)) &= 1429 \\ NPV(0.1) &= 966 & \sigma(NPV(0.1)) &= 382 \\ IRR &= 24.6\% & \sigma(IRR) &= 3.4\% \end{aligned} \tag{4.8}$$

This policy performs approximately the same as the base case. Although being more optimistic, this policy is no more risky. Clearly, optimisation under deterministic reserves is significantly worse than stochastic optimisation. The reason being that a number of reservoirs are potentially large, but have a significant probability of being empty. The best policy is to develop these reservoirs, since they may be quickly abandonned if found to be empty. However, under deterministic reserves these reservoirs are considered to be small (since their average size is small) and might not be developed. This shows the error of assuming the size of the reservoirs to be known - one cannot account for changes in policy once better information is gained.

4.4.5 The worth of knowing reserves more precisely

As outlined previously (see sec. 4.2.3) one may wish to evaluate the worth of further exploration of a given reservoir. Given as cluster 9 potentially contains a large amount of oil, an investigation on this cluster was considered in which the presence of oil (or not) is determined for the entire cluster. The investigation was assumed to give no further information about the amount of oil in the cluster. This cluster contains 5 reservoirs, each of which have a different probability of containing oil. To combine these reservoirs, an average probability for the presence of oil was found, weighted by the expected size of the reserves if each reservoir is present.

A pair of optimisations were made, one for which the cluster was present, and one for which it was not. The optimisations were performed by taking the optimal policy found above and performing a quench in the two scenarios, where only the hub capacity and location were allowed to change. This leads to a new set of expected financial measures of

$$\begin{array}{ll} NPV(0) = 8818 & \sigma(NPV(0)) = 1768 \\ NPV(0.1) = 1488 & \sigma(NPV(0.1)) = 517 \\ IRR = 25.3\% & \sigma(IRR) = 3.3\% \end{array} \tag{4.9}$$

The value of $NPV(0.1)$ is slightly improved on the optimal case. However, $\sigma(NPV(0.1))$ is less than in the optimal case, indicating that the extra information reduces the uncertainty in the project. Whether these improvements are sufficiently large to justify such an exploration is a matter for BP. What is of interest is the confirmation of

the obvious - that extra information is of value, and that this can be demonstrated in a commercial setting.

4.4.6 Penalty for zero reserves

If a reservoir contains no oil, then some penalty should be paid, since wells will need to be drilled to establish this. Since this cost officially falls under the exploration budget, it has not been included in the original model. However, to include a penalty for a reservoir being empty allows for a more fair comparison of differing policies. We consider here the penalty to be the cost of one appraisal well, paid when the production wells would have been drilled. The policy found is significantly less optimistic than the optimal policy, but still involves drilling a large number of reservoirs (as illustrated in figure 4.9). The financial measures under this penalty for this policy and the base case are

$$\begin{array}{lll} \text{Optimised policy} & NPV(0) = 7422 & \sigma(NPV(0)) = 1930 \\ & NPV(0.1) = 1196 & \sigma(NPV(0.1)) = 549 \\ & IRR = 24.2\% & \sigma(IRR) = 3.7\% \end{array} \tag{4.10}$$

$$\begin{array}{lll} \text{Base case} & NPV(0) = 5070 & \sigma(NPV(0)) = 1396 \\ & NPV(0.1) = 861 & \sigma(NPV(0.1)) = 412 \\ & IRR = 24.2\% & \sigma(IRR) = 3.9\% \end{array} \tag{4.11}$$

Noticeably, the difference between the expected NPV of these two policies is not as large as between the base case and optimisation with no penalty, suggesting that the base case is robust with respect to penalties for empty reservoirs.

4.5 Conclusions

The aim of this study has been to demonstrate the applicability of stochastic optimisation techniques, and in particular stochastic annealing, to a commercial problem where uncertainty plays a key role. This has been demonstrated, with the optimisation producing a policy with a 50% improvement in $NPV(0.1)$.

Most of the policies found by optimisation fall into the class of a "big bang", which means that production is very high for a short number of years, and falls off

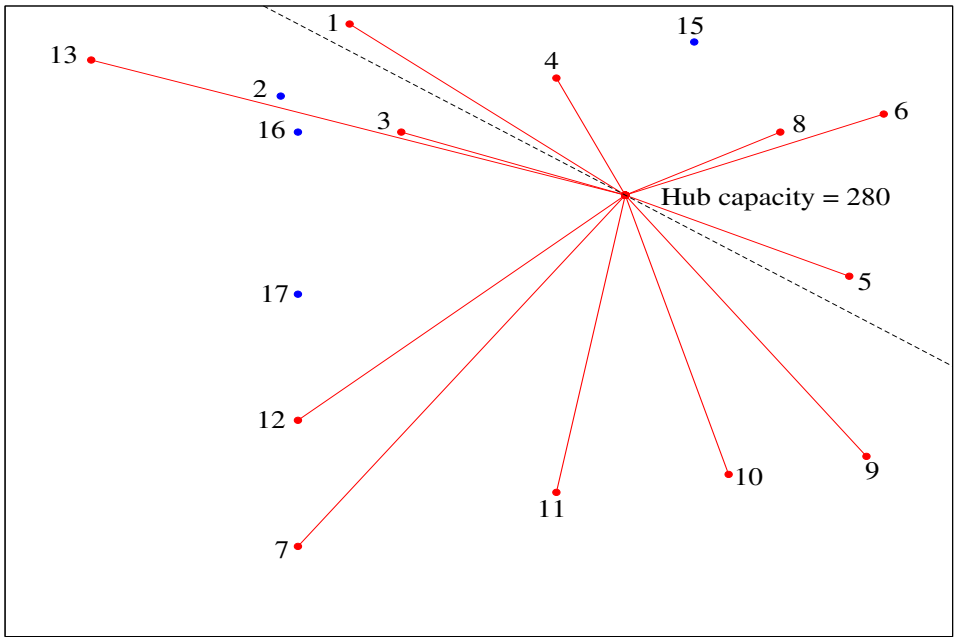

Figure 4.9: The configuration of the policy chosen by stochastic annealing when a penalty of the cost of one appraisal well is given when a reservoir to be developed is empty. This policy develops 12 clusters and is significantly less optimistic than the optimal policy for zero penalty (many fewer reservoirs are drilled in each cluster). The expected NPV of this policy is $NPV(0.1) = 1196$.

rapidly. This is due to the high (but not unrealistic) discount factor. The base case is a more gentle strategy, presumably for staffing or logistics reasons which are not included in the model.

We have seen that the uncertainty has a significant effect on the optimal policy for the development of the oil field and hence uncertainty needs to be taken into account in oil field models. This not only implies the inclusion of reserves uncertainty in all future models, but also the development of models which include other areas of uncertainty. Since price uncertainty is a dominant factor [66] and price fluctuations are future events, the natural approach is to develop oil field models that are flexible, and can be altered as the oil price changes. Stochastic annealing is a technique well placed to deal with such developments.

Chapter 5

Diffusion limited aggregation

A straight line may be the shortest distance between two points,
but it is by no means the most interesting. – Dr. Who

Diffusion limited aggregation (DLA) is an extensively studied model of diffusion limited growth since it is a simple model which appears to capture the essential features of many different physical phenomena. It is also of interest since the fractals generated have evaded a complete understanding for many years. This chapter mainly deals with the topic of noise reduction when DLA clusters are grown off-lattice. This is a new development that permits one to measure the properties of DLA clusters much more accurately, and ties in with a renormalisation scheme. We begin by outlining previous developments on DLA followed by the introduction of noise reduction off-lattice. We then consider the correction to scaling exponents of the clusters.

5.1 Introduction

5.1.1 Why DLA is interesting

Witten and Sander [114] noted that a number of systems of aggregating particles have a very similar appearance. They proposed a model of these systems which is called diffusion limited aggregation (DLA). Since then, many systems have been seen to exhibit diffusion-limited growth [97]. Examples include electrodeposition [17], viscous fingering [82], bacterial colonies [35] and snow crystals [57]. They all share the common appearance of growing fingers, which split and compete (see fig.

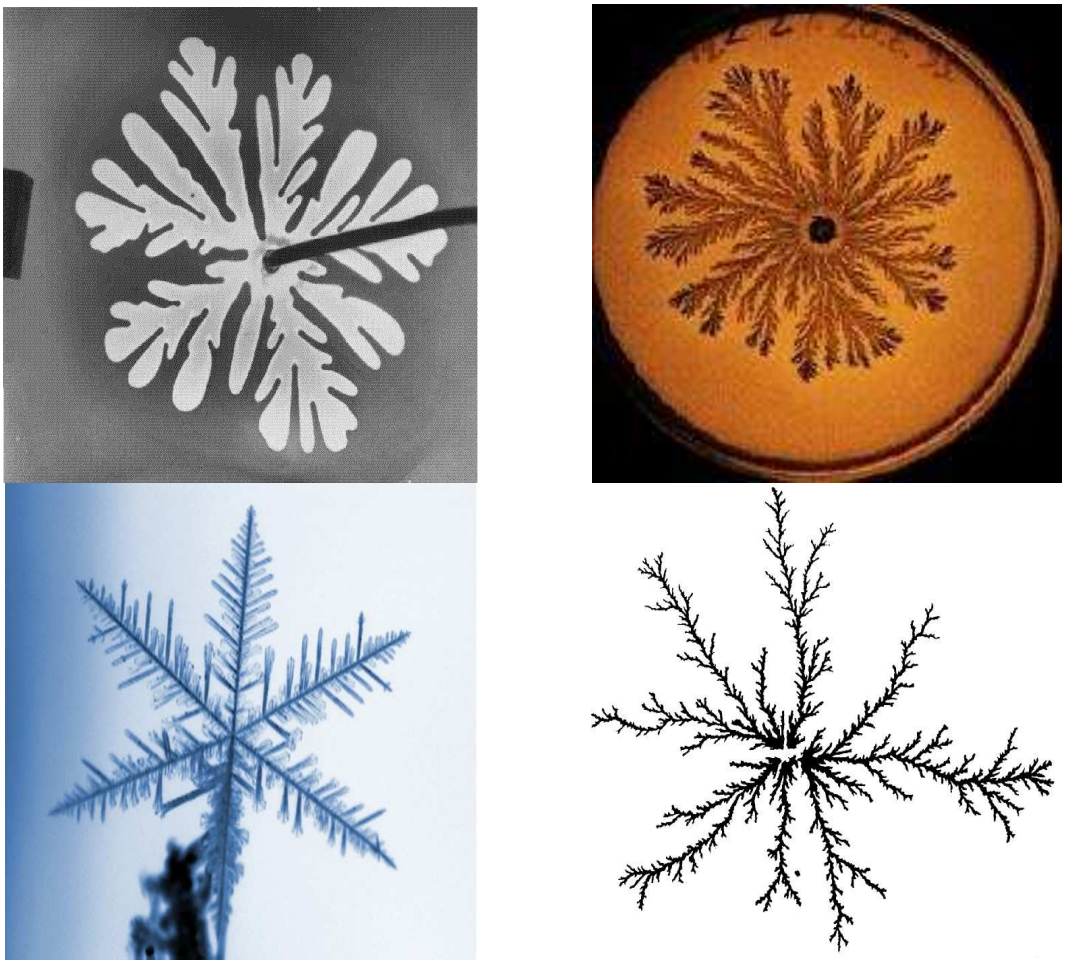

Figure 5.1: Examples of diffusion limited growth in real systems. These are a viscous fingering pattern [97] (top-left), a bacterial colony [97] (top-right), a snow crystal [62] (bottom-left) and an electrodeposit [71] (bottom-right).

5.1).

Since DLA is an idealised system, it does not precisely replicate the behaviour of the real systems. For example, viscous fingers grown in a channel are not always fractal [110], unlike DLA grown in cylindrical geometry [32]. DLA is a simple algorithm which produces random fractals and has evaded a complete understanding for many years. Hence, DLA provides an interesting test for any theories which aim to explain the growth of fractals.

5.1.2 The original work

We will consider diffusion-limited growth where particles are allowed to diffuse freely until they contact the cluster, at which point they are irreversibly stuck. The cluster

may be grown from a seed of any shape, but we will consider the seed to be a single particle. If $\rho(\mathbf{r}, t)$ is the density of diffusing particles, this will obey the diffusion equation

$$\frac{\partial \rho}{\partial t} = K_{\text{diff}} \nabla^2 \rho \tag{5.1}$$

where K_{diff} is the diffusion constant. We consider the the growth of the boundary to be slow and hence $\frac{\partial \rho}{\partial t} = 0$. Thus the growth is governed by Laplace's equation

$$\nabla^2 \rho = 0. \tag{5.2}$$

The growth of this aggregate can be considered from the solution of an electrostatic problem. Since the cluster is considered to be a perfect absorber, we have $\rho = 0$ on the cluster. We also have to chose the appropriate boundary condition at infinity, which is $\rho \to \ln(r) + \text{const}$ for clusters grown in 2 dimensions, and $\rho \to \rho_\infty$, a constant, in higher dimensions.

Description of the growth algorithm

Diffusion on a computer is modelled by a number of particles performing a random walk, mirroring the Brownian motion of particles in a gas. DLA is modelled [114] by allowing particles to randomly walk from "infinity", one at a time, until they contact the cluster, at which point they are irreversibly stuck. The alternate situation, where we have a large number of random walkers, which are allowed to stick to each other and where the clusters remain mobile is called cluster-cluster aggregation (CCA, or DLCA) [75, 54, 107]. CCA clusters have different properties from DLA clusters - and we will restrict our attention to DLA.

One need not launch each random walker from infinity in practice, but from a random location on a circle which encloses the cluster. The original model is on a square lattice, so at each time step, the random walker moves a square in one of four directions, illustrated in figure 5.2. This also shows the kill radius, R_{kill} beyond which diffusing particles are killed. This must be chosen to be sufficiently large so as not to affect the dynamics. An example of an on-lattice DLA cluster is shown in figure 5.3. This cluster was grown to 5,000 particles, with the darker shading signifying a later time of deposition on the cluster.

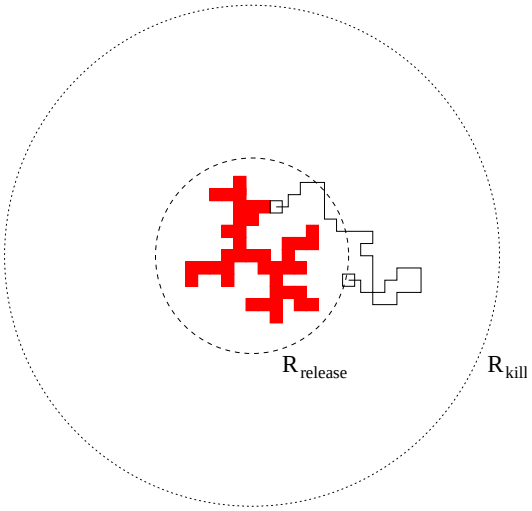

Figure 5.2: The computational method for DLA on-lattice. A diffusing particle is allowed to randomly walk from the release radius, until it contacts the cluster, at which point it is irreversibly stuck. If the particle reaches the kill radius, it is deleted, and a new particle started from the release radius.

These seem to be fractals

The clusters that are grown using this model appear to be fractal. We will reconsider the definition of a fractal later (see section 5.1.3). For now we will take the definition to be

$$M(r) \propto r^D \tag{5.3}$$

where $M(r)$ is the mass of the cluster within radius r of the seed, and D is the dimension of the cluster. Objects with non-integer values of D we shall consider to be fractals. For DLA clusters grown off-lattice we find that $D \simeq 1.71$. Witten and Sander [114] found a similar value of the fractal dimension for moderate size clusters grown on-lattice, but this value is not correct asymptotically for DLA grown on-lattice as we shall see later.

The distinctive characteristics of DLA and DLA-like growth are extended branches and deep fjords. This is due to the Mullins-Sekerka instability [80] and the screening of fjords. Consider that we allow a circle to grow (without noise) at a rate given by a diffusive field. The circle will grow uniformly and remain as a circle. However, this growth is not stable and if we consider small bump on the surface of the circle, this bump is expected to grow faster than the rest of the circle [80]. This

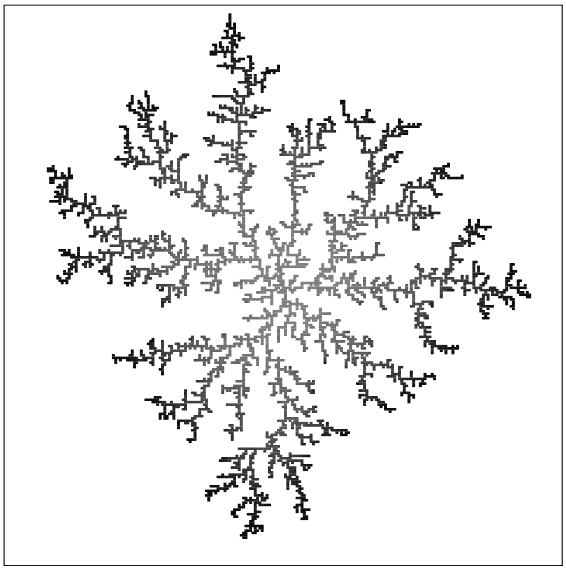

Figure 5.3: A small DLA cluster grown on-lattice with $N = 5,000$. The darker shading signifies that point on the cluster being grown at a later time.

is the Mullins-Sekerka instability which means that branches will grow faster than fjords, and can lead the tips of the branches to split. A random walker may stick to the side of a branch. Under favourable conditions this protrusion may become a side-branch and eventually dominate the growth. The screening of fjords is caused by the same effect in reverse. If we have a circle with a small indent in it, then this indent will grow more slowly than the surrounding circle, leading to the screening of fjords.

The screening of fjords and tip-splitting are thought to be [97] the dominant processes in creating the highly branched fractal structure of DLA.

5.1.3 Aside on fractals

Unfortunately, no watertight definition of a fractal exists ([33], page 11). Many tentative definitions are available, but none appear to perfectly encompass all the objects which are considered to be fractal. To give a qualitative view, 2 definitions are offered, both of which classify DLA clusters as fractals.

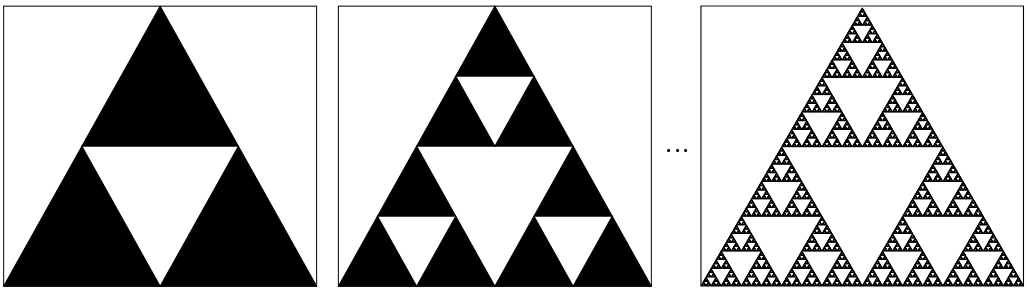

Figure 5.4: Sierpinski's gasket. At each stage the triangles are replaced with 3 smaller triangles (left). When repeated this leads to a fractal structure (right) with fractal dimension $D \simeq 1.58$.

Self-similarity

"A fractal is a shape made of parts similar to the whole in some way" ([67], page 17). That is, a fractal appears similar to itself on all scales. An example is the Sierpinski gasket, a fractal of dimension $D = \frac{\ln 3}{\ln 2} \simeq 1.58$ [33]. The gasket is created from a set of 3 triangles as in fig. 5.4 (left). Each of these 3 triangles are replaced by 3 further triangles, scaled down by a factor $\frac{1}{2}$. This procedure is repeated ad. infinitum to create the self-similar fractal (fig. 5.4 (right)).

DLA clusters appear to have this property of self-similarity. Each branch has many sub-branches which in turn have their own sub-branches etc. DLA is, however, only self similar for averaged quantities (such as the average distance that a branch grows before branching) and is therefore described as *statistically self-similar*.

Density correlations

A fractal is a shape whose cumulative two point correlation function follows

$$C_2(r) \propto r^D \tag{5.4}$$

where D is the (non-integer) fractal dimension, and $C_2(r)$ is defined by

$$C_2(r) = \frac{1}{N} \sum_{\mathbf{r}'} M(\mathbf{r}', r) \tag{5.5}$$

where N is the number of particles in the cluster, $\mathbf{r}'$ is the position of a particle in the cluster, the sum is over all particles in the cluster, and $M(\mathbf{r}', r)$ is the number of particles within a radius r of $\mathbf{r}'$.

This definition of a fractal is similar to one proposed above (eq. 5.3), and DLA satisfies both. We interpret this definition as meaning that DLA clusters are

riddled with holes on all length scales so that they fail to fill the space in which they are grown.

5.1.4 The penetration depth

Since DLA is characterised by deep fjords (and holes of all sizes) some growth must be happening at the edges of the cluster, and some closer to the seed. Plischke and Racz [65] introduced the penetration depth (or active zone) of a cluster, ξ, which is defined as the standard deviation of $P(r)$, the probability density of a diffusing particle first contacting the cluster at radius r from the seed. If DLA is a fractal, we expect that $\frac{\xi}{R_{dep}}$ will tend toward a constant for large clusters, where R_{dep} is the average radius of deposition. Plischke and Racz [65] found that ξ grows more slowly than R_{dep}. This is troublesome for DLA, since it would mean that deposition would be occurring within an increasingly thin fraction of the cluster. Since the measured penetration depth also gives the size of the gaps between arms [104], this would lead us to expect the cluster to be compact (i.e. D=2).

5.1.5 Large clusters

The problem of the scaling of the penetration depth was answered in 2 different ways. Meakin & Sander [76] grew a number of clusters off-lattice, and Ball & Brady [5] developed a technique for growing large DLA clusters.

Meakin & Sander [76] measured the penetration depth of a large number of clusters grown off-lattice. For small clusters the penetration depth appeared to have a smaller exponent than the radius. However, larger clusters showed an increasing exponent and they concluded that asymptotically ξ and R scale at the same rate.

Ball & Brady [5] suggested a computational trick to allow large clusters to be grown efficiently. One covers the entire cluster with a grid whose square sizes are large. If we have a square which contains no sites which border the cluster, then that square is said to be empty. Each occupied square is sub-divided into 4 smaller squares and the procedure is repeated down to some predefined depth. If a diffusing particle happens to be in an empty square (of whatever size), then it may freely move to the edge of the square; its position being given by the solution of Laplace's equation in an empty square. This allows the particle to take large steps

when it is not near the cluster, and provides a much more efficient algorithm than the naïve method. There are other implementations of this scheme, which do not involve finding the solution of Laplace's equation on a square. Additional efficiency savings can also be made by combining this with other techniques [73, 97].

This method allowed Ball and Brady to grow much larger clusters than were previously available. They showed that DLA clusters show a bias due to the lattice. They also found that the penetration depth does scale with the radius, and that the different scaling of the penetration depth, as seen by Plischke and Racz [65], was due to their growing only small clusters. This has since been confirmed by Ossadnik for very large off-lattice clusters [87].

At this point, progress in DLA diverged. On one hand, researchers considered the properties of off-lattice DLA by growing as large clusters as possible. This has lead to a few interesting questions about the nature of isotropic DLA, which we will deal with later. On the other hand, on-lattice DLA has seen the introduction of noise reduction and the use of analytical techniques, which is considered now.

5.1.6 Noise reduction on-lattice and lattice anisotropy

One crucial development for the understanding of DLA is noise reduction [109]. This allows us to probe the behaviour of large DLA clusters without growing large clusters. This development answered a number of questions about the role of the lattice in DLA.

The scheme

In normal DLA, a site is added to the cluster when a random walker has landed on it. For noise reduced DLA, when a random walker lands on a site adjacent to the cluster a counter on that site is increased by 1, the random walker is erased, and a new random walker is launched. Once that counter has reached m, the site grows and becomes part of the cluster. When the site is grown, all the new neighbouring sites begin with a counter set to zero. Thus, $m = 1$ corresponds to ordinary DLA, and large m corresponds to large noise reduction.

We still have the discreteness of only being able to grow a whole site at a time. However, the requirement that multiple walkers land on a site is equivalent

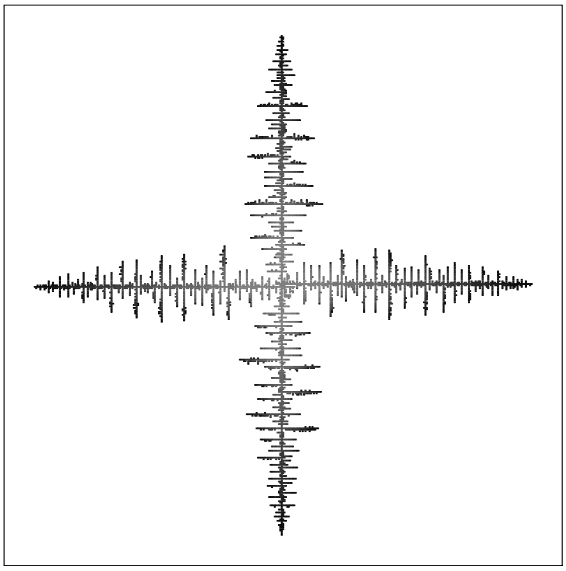

Figure 5.5: A noise reduced DLA cluster, on-lattice, with $N = 5,000$ and $m = 100$. Darker shading indicates that a given cluster particle was grown later. Although the bias of the lattice was not obvious in the normal DLA cluster (fig. 5.3) it is clearly seen at high noise reduction.

to more accurately sampling the Laplacian field before growth is allowed.

The problem - this does not appear to be the same as "normal" DLA

A typical DLA cluster with noise reduction ($m = 100$ and $N = 5,000$) is shown in fig. 5.5. The cluster clearly displays a growth bias caused by the presence of the lattice. This bias is severe, and noise reduced DLA clusters (grown on-lattice) do not show the expected fractal dimension of 1.71, but instead one finds $D \simeq 1.5$ for these clusters. Even large DLA clusters with $m = 1$ do not appear to have this needle-like form, although they do show a lattice bias. One may wonder whether noise-reduced DLA is the same as ordinary DLA, or is showing some particular feature of the noise reduction.

Lattice anisotropy

Eckmann et al. [31] were able to characterise the shape of a noise reduced DLA cluster, giving an envelope in which the cluster is contained (see fig. 5.6) using a conformal map technique. The shape of these envelopes were found to be independent of m, and it was concluded that on-lattice DLA has a fractal dimension D

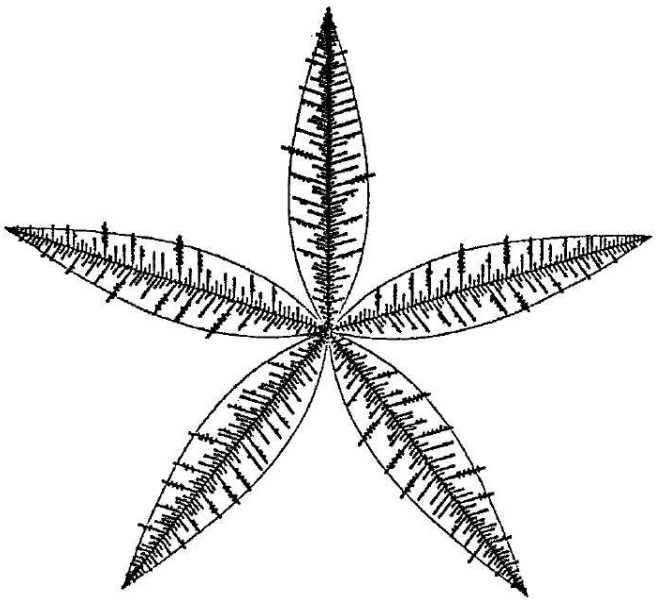

Figure 5.6: A noise-reduced DLA cluster 5-fold anisotropy. It is contained within an envelope which is set for all clusters. This figure is taken from [31].

closer to 1.5, and that the value of 1.71 is only a transient for small clusters. The cluster size required to verify this for $m = 1$ directly is huge, $O(10^7)$ particles [77].

Noise reduced DLA clusters lead to an interesting comparison with snow crystals (cf. figs. 5.1, 5.5). Snow crystals clearly show a six-fold lattice bias, with a shape that resembles on-lattice clusters. This is believed to be due to an anisotropy introduced by the way water molecules form into crystals [83, 84].

5.1.7 Controversy

With consideration of the penetration depth, we have the concept of an active-zone where growth is occurring, and a dead-zone where no further growth will take place. Mandelbrot et al. [69, 70] considered various moments of DLA clusters, focusing on the values they take within the dead-zone of a cluster. They claimed that their data was consistent with an infinite drift scenario [68] (which would mean that DLA clusters are not fractal). However, they were not able to rule out a simple scaling behaviour.

Certain properties of DLA clusters are still open to question, and need to be

clarified. Noise reduction on-lattice has proved to be a very important technique in the understanding of DLA. Thus, we might expect that noise reduction off-lattice will prove equally as useful.

5.2 Off-lattice noise reduced diffusion limited aggregation

Noise reduction has been an important development in the study of DLA. It allows one to grow clusters which have the same properties as very large clusters of 'normal' DLA. This has led to the study of lattice effects and the development of renormalisation schemes (see sec. 6.1.2). However, the lattice effects are severe, and on-lattice DLA is not an accurate representation of diffusion limited growth [31]. Most of the recent work on DLA has been off-lattice [97]. Hence, a noise reduction technique off-lattice would be most useful, as it would allow further developments in the understanding of DLA.

In this section, a technique is introduced for noise reduction off-lattice, and it is argued that this is a good choice of algorithm. We consider its relation to the HL method [44] for growing noise reduced clusters, and acknowledge that both techniques have particular benefits.

5.2.1 Definition of method

We place a seed particle at the centre of the co-ordinate system. Another particle is allowed to diffuse freely from a circle surrounding the seed in the usual manner. Once the diffusing particle contacts the cluster, it is moved onto the cluster particle, reducing the distance between the particle's centres by a factor of A (the noise reduction). The diffusing particle and the cluster particle overlap, and what is added to the cluster is a crescent. Figure 5.7 illustrates a diffusing particle hitting the seed, for normal DLA and for noise reduced DLA, showing not only the particle picture, but also the capture boundary. The capture boundary is the boundary which is 2 particle radii from the centre of cluster particles. If the centre of a diffusing particle crosses the capture boundary of the cluster, the particle will be in contact with the cluster. Normal DLA is characterised by adding bumps of aspect ratio $\frac{1}{2}$ to the

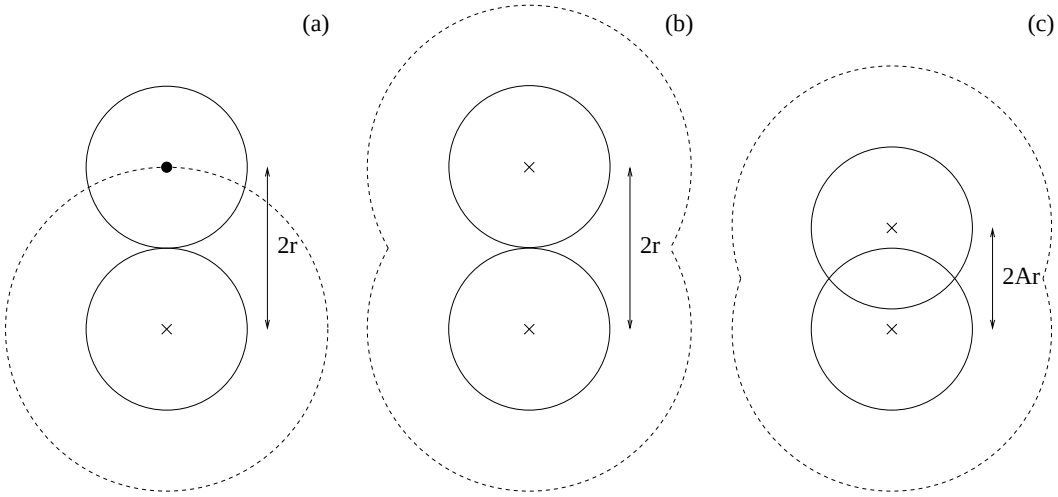

Figure 5.7: The noise-reduced DLA algorithm. In normal DLA, a particle of radius r diffuses until it contacts the seed (a). The capture boundary for the seed is also shown, the particle being in contact with the cluster when its centre is on the capture boundary. In normal DLA, the particle is deposited at the point of first contacts the cluster (b). This results in a bump of aspect ratio $\frac{1}{2}$ being added to the capture boundary. In noise reduced DLA, (c), the diffusing particle is moved onto the cluster by a factor of A. This means that a bump of aspect ratio A is added to the cluster, and a bump of aspect ratio $\frac{A}{2}$ added to the capture boundary.

capture boundary. Noise reduced DLA is characterised by adding bumps of aspect ratio $\frac{A}{2}$ to the capture boundary, and aspect ratio A in the standard picture of the cluster. It is clear that $A = 1$ corresponds to no noise reduction, and we hope that A small will correspond to highly noise reduced clusters in the same way that $\frac{1}{m}$ small does for on-lattice noise reduction.

5.2.2 Justification that it is DLA - conceptually

This technique should produce clusters that are noise reduced versions of normal DLA clusters. The addition of shallow bumps to the cluster should reduce the shot noise of adding new particles. In the case of noise reduction on-lattice, one had to contact a site m times before it was grown. In the off-lattice algorithm, if one contacts the cluster $\frac{1}{A}$ times in the same place, it will produce a protrusion which reaches out 1 particle diameter.

Fig. 5.8 shows a number of DLA clusters grown using off-lattice noise reduction. A highly noise reduced cluster with $N = 10^4$ and $A = 0.03$ (a) is shown. Compared to a cluster with $N = 10^4$ and $A = 1$ (b), the noise reduced cluster

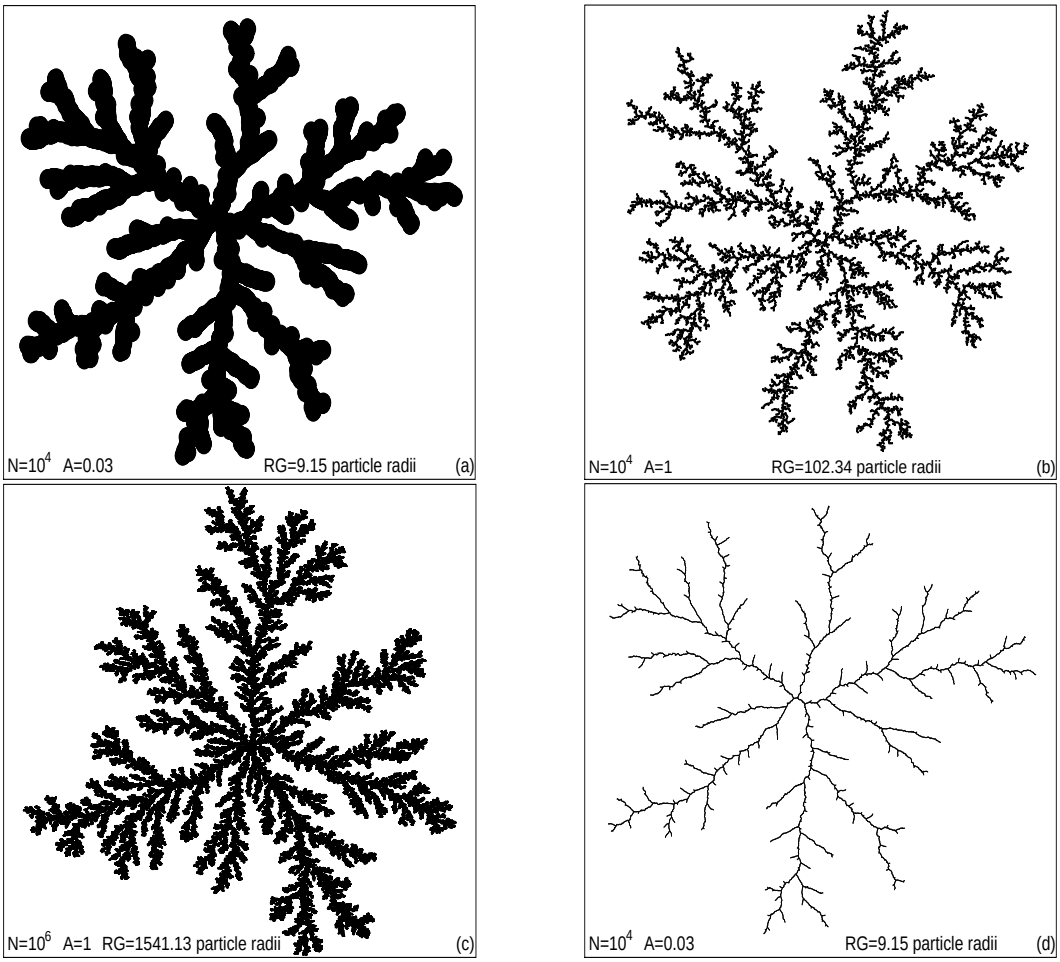

Figure 5.8: Some DLA clusters grown using this noise reduction method. The first is a cluster with 10^4 particles and $A = 0.03$ (a). It will be noted that this cluster appears more mature than the cluster with 10^4 particles and $A = 1$ (b). The noise reduced cluster has the same properties as the cluster with 10^6 particles and $A = 1$ (c). The noise reduced cluster is also shown where lines are drawn between adjacent particles to highlight the fine structure of the cluster (d).

appears to be much more mature, which we quantify below. In fact, it has the same properties as a normal DLA cluster with $N = 10^6$ (c). The noise reduced cluster is also shown with lines joining adjacent particles in the cluster, to emphasise the fine structure of the cluster (d).

5.2.3 Fractal dimension

To further test whether the noise reduction method is correctly simulating DLA, the fractal dimension of noise reduced DLA clusters is measured. The fractal dimension

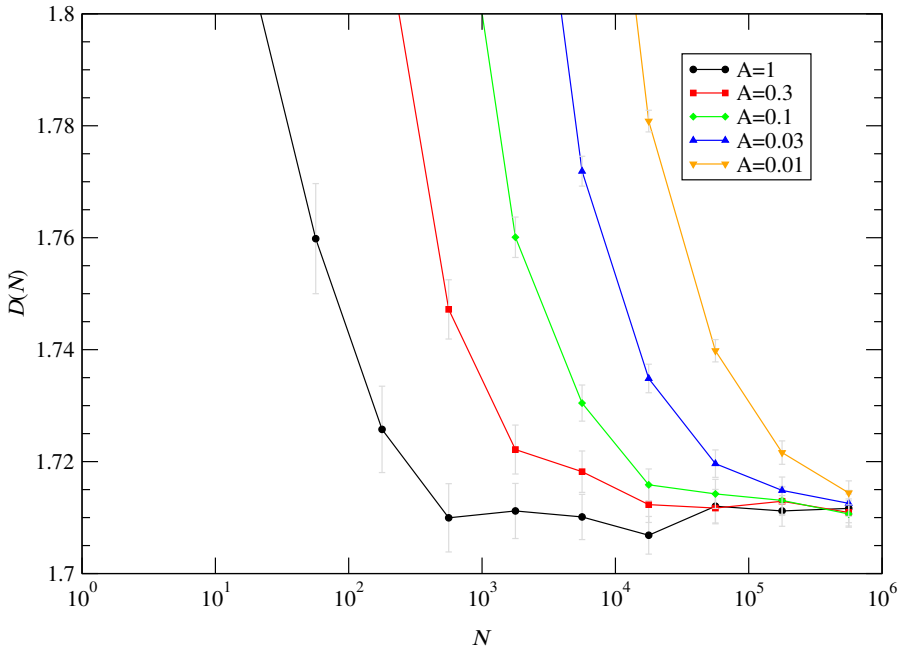

Figure 5.9: The fractal dimension of DLA, measured from the local slope of R_{dep} against N. The clusters for all values of A are converging to the same value, indicating that noise reduction off-lattice is correctly simulating DLA. The measured asymptotic value is $D = 1.712 \pm 0.003$.

is calculated from the local slope of the average radius of deposition according to

$$D = \frac{\ln(N_2) - \ln(N_1)}{\ln(R_{dep}(N_2)) - \ln(R_{dep}(N_1))}. \tag{5.6}$$

Details on how the clusters were grown are given in sec. 5.3.1. The fractal dimension of DLA clusters is plotted in figure 5.9. This shows that the clusters for all values of the noise reduction are converging to the same value, and we measure the dimension to be $D = 1.712 \pm 0.003$, consistent with previous measurements [26]. Note that the measured dimension for $A = 1$ appears to be slowly drifting upwards. This hints that the clusters $A = 1$ may not be fully mature. This measurement not only illustrates that noise reduced DLA is accurately simulating DLA, but that it is possible to more accurately measure the asymptotic properties of DLA than previously.

5.2.4 Relative penetration depth

The scaling of the penetration depth has been a topic of some controversy (sec. 5.1.4). Figure 5.10 shows data for the relative penetration depth, $\frac{\xi}{R_{dep}}$, of DLA clusters

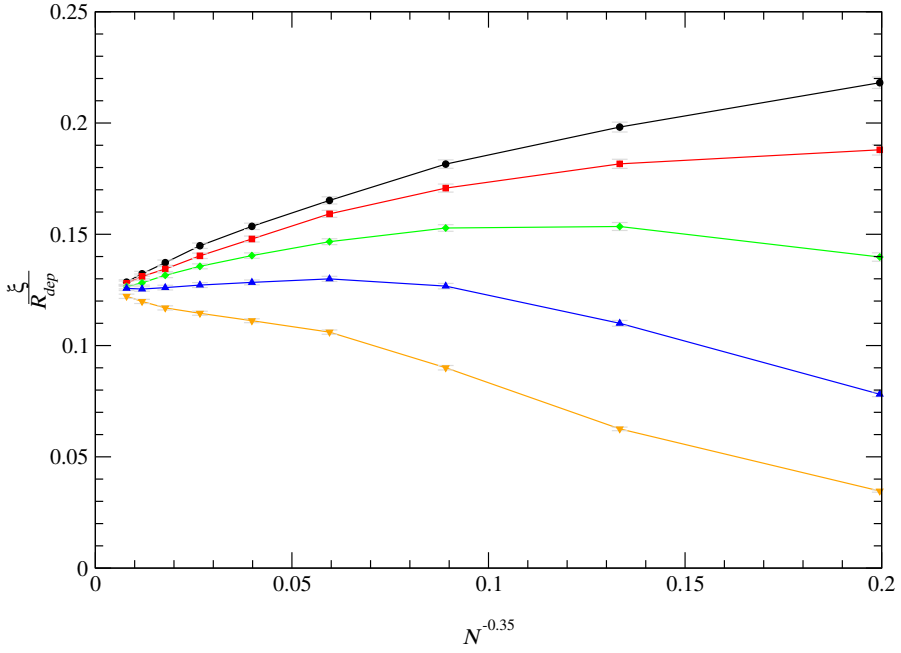

Figure 5.10: The relative penetration depth $\frac{\xi}{R_{dep}}$ as a function of $N^{-0.35}$. 0.35 is chosen as this is the measured correction to scaling exponent for $\frac{\xi}{R_{dep}}$ (see sec. 5.5). The data are converging to a single, finite value of $\frac{\xi}{R_{dep}}$ as $N \to \infty$.

grown using off-lattice noise reduction. Note that all the data are approaching the same asymptotic value from different points. This allows us to pin down the asymptotic value very accurately, and it is found that $\frac{\xi}{R_{dep}} = 0.125 \pm 0.001$.

5.3 Relation to the HL method

This technique for off-lattice noise reduction in DLA is not the first of its kind. Hastings and Levitov [44] proposed a technique using conformal maps to model Laplacian growth. A number of papers [24, 25, 26, 27] have recently appeared using the HL method to grow DLA clusters. Unfortunately, the computer time required by this method is $O(N^2)$, meaning that only (relatively) small clusters may be grown.

The method works by a conformal mapping of the unit circle to the perimeter of a DLA cluster. This is achieved by the repeated composition of the following map

$$\phi_{\lambda,a}(z) = z^{1-a}\left((1+\lambda)\frac{1+z}{2z}\left(1+z+\sqrt{1+z^2-2z\frac{1-\lambda}{1+\lambda}}\right)-1\right)^a \qquad (5.7)$$

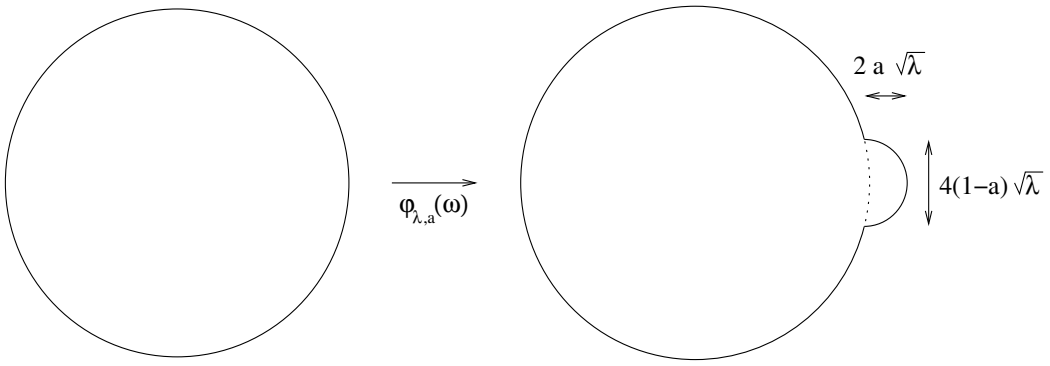

Figure 5.11: The primitive map for the HL method, where the unit circle is mapped to the unit circle with a bump. The size of this bump is given for $\lambda \ll 1$. By repeated composition of this primitive map, one is able to map the exterior of the unit circle to the exterior of a DLA cluster.

where λ and a are parameters, which maps the unit circle to the unit circle with a bump (see fig. 5.11). The noise reduction parameter, a, controls the shape of the bump, and the parameter λ controls the size of the bump which must be chosen to ensure that the bump, when mapped, has a fixed size. If we compose these maps according to the following form

$$F_N(z) = F_{N-1}(e^{i\theta_n}\phi(ze^{-i\theta_n})) \tag{5.8}$$

where F_N is the map for a cluster of N particles and θ_N is a randomly chosen angle at which the new bump is deposited, then this algorithm produces clusters which show the correct scaling for DLA [44].

Theoretical relation

Since off-lattice noise reduction and the HL method are attempting to model the same situation, we expect there to be a relationship between the two noise reduction parameters, A and a. The HL method maps the unit circle to the capture boundary of a DLA cluster, and for $\lambda \ll 1$ one estimates the aspect ratio of a bump in the HL method to be $\frac{a}{2(1-a)}$. Hence, $a = \frac{1}{2}$ should correspond to normal DLA. One expects the following relation between the two methods of noise reduction

$$A = \frac{a}{1-a}. \tag{5.9}$$

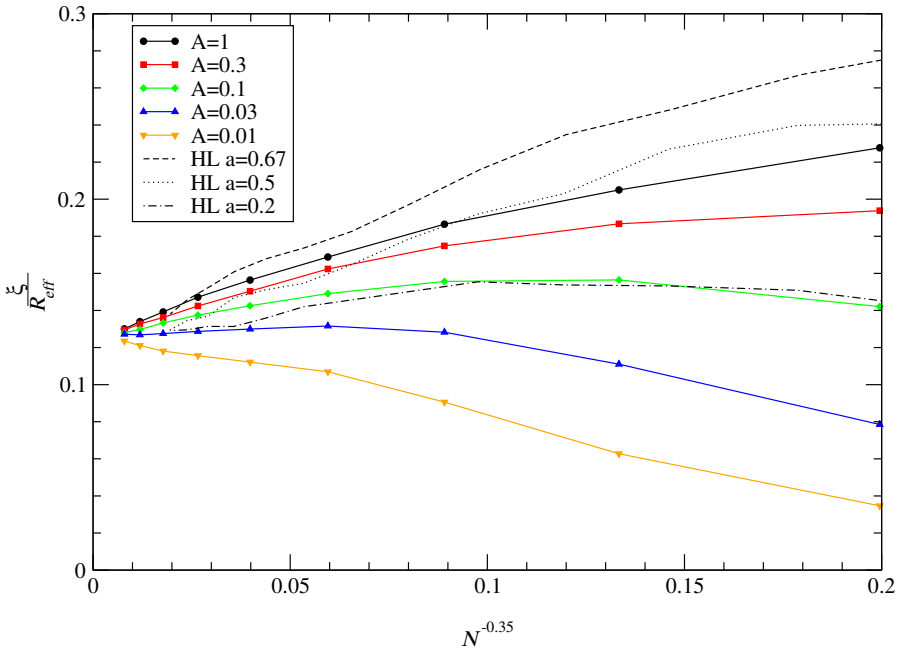

Figure 5.12: The penetration depth for HL clusters and noise reduced DLA clusters. The data are well matched for many cluster sizes, but the HL appears to deviate from the noise reduced DLA data for large clusters. This may be due to the problem of fjord filling.

Comparison of the properties of the clusters

We now compare the properties of DLA clusters grown using the HL method with noise reduced DLA clusters (the data from the HL method is courtesy of Len Sander [96]). Figure 5.12 shows a comparison for the relative penetration depth (in this case $\frac{\xi}{R_{eff}}$) where R_{eff} is the effective, or Laplace, radius of the cluster (see sec. 5.4.3). The data appear well matched for many cluster sizes, demonstrating that this noise reduction technique is correctly simulating DLA. $a = 0.2$ should correspond to $A = 0.25$. However, the penetration depth for the HL data is significantly smaller than the data for $A = 0.1$ for large cluster sizes. This is expected, since "fjord filling" events [25] have been observed for a small, and such events would artificially reduce the penetration depth. This drift is also observed for both $a = 0.5$ and $a = 0.67$, which should correspond to $A = 1$ and $A = 2$ respectively. This suggests that fjord filling is a problem for all values of a, and the use of $a = \frac{2}{3}$ [25, 27, 28] does not avoid this problem as others have hoped.

The problem of fjord filling is not intrinsic to the HL method. A value for λ must be chosen to ensure that bumps, when mapped, are all the same size. The method for estimating this size can be inaccurate, and if the wrong size is chosen, this can lead to a fjord being filled. However, it is quite possible to choose a more accurate method for this estimation, and one such method has been suggested by Stepanov & Levitov [106]. Their method chooses to reject all bumps that are outside of a threshold size when mapped. A much more satisfying approach would be to choose λ accurately so that all bumps are of identical size. Given that fjord filling affects the HL method for all values of a, new methods for choosing bump sizes need to be developed.

The usefulness of both methods

The advantage of the HL method is that it allows one to *exactly* extract the harmonic measure of a cluster [27]. The harmonic measure is the probability distribution $P(s)ds$ that a diffusing particle will contact the cluster between s and $s+ds$ on the boundary (see [33], page 92). It has recently been claimed [49] that previous attempts to calculate the harmonic measure for DLA were poorly converged, since it is difficult to measure the growth probability for sites where the growth probability is extremely small. This means that the HL method is expected to the the technique of choice for evaluating the harmonic measure of DLA. If, however, the desire is to grow clusters which are fully mature, then off-lattice noise reduction should be used. This technique has the benefit of being fast and applicable in any dimension.

5.3.1 Computational details and collaboration

A number of noise reduced DLA clusters were grown using off-lattice noise reduction in 2 and 3 dimensions. For both 1,000 clusters were grown to 10^6 particles with 5 values of A ($A = 1, 0.3, 0.1, 0.03, 0.01$). The growth was stopped every $10^{\frac{i}{2}}$ particles (i integer) and a number of probe particles were fired at the cluster. Probe particles are allowed to diffuse normally, but when they contact the cluster, their position is recorded and they are deleted. Thus, they are used to probe the cluster at fixed N. In 2D 10^4 probe particles were fired each time, and in 3D 10^5 were used.

Once the locations of all the probe particles had been noted, the data was

re-centred to the centre of charge so that the monopole moment is zero, $P_1 = 0$ (see sec. 5.4.3). The only exception to this is the extremal cluster radius, R_{ext}, which was measured with respect to the seed, as this provided more stable data.

The simulations were performed off-lattice using the Ball & Brady [5] heirarchical maps technique, using a code developed by T. Rage [90]. On the lowest level the particles made steps of 1 particle radius in size. If a particle landed on the cluster, it was "backed out" to its point of first contact. This work was carried out in collaboration with Ellak Somfai and Len Sander. The simulations of Ellak Somfai [103] (in 2 dimensions) used a much smaller minimum step size, and his results are found to be in good agreement with the ones presented here. All the data for the clusters grown using the HL method are courtesy of Len Sander [96].

5.4 Correction to scaling exponents

5.4.1 What is a correction to scaling exponent?

Let us suppose that we have some quantity, Q, which tends toward a particular value, Q_∞, for large clusters. Q is said to have a single correction to scaling if its value for a cluster of size N is

$$Q = Q_\infty + CN^{-\nu} \tag{5.10}$$

where C and ν are constants and ν is called the correction to scaling exponent.

5.4.2 Why single slowest correction to scaling exponent is important

It has recently been demonstrated [104] that the penetration depth ξ is given by

$$\xi^2 = \frac{1}{2}\sum_{i=1}^{\infty} |B_{-i}|^2 \tag{5.11}$$

where B_{-i} is the Laurent coefficient of z^{-i} of the conformal map from the unit circle to a DLA cluster. Since these coefficients completely characterise the cluster, no quantity can exhibit a slower correction to scaling than that shown by the penetration depth, unless contributions from different coefficients cancel fortuitously.

Further, it was conjectured [104] that all the properties of DLA will show this same correction to scaling exponent.

5.4.3 Definition of multipole moments

To be capable of testing whether DLA has a single correction to scaling, we need a set of measurements that completely characterise the cluster. Such a set of measurements is provided by the multipole moments [46]. The analogy between Laplacian growth and electrostatics (see sec. 5.1.2) allows us to use them to analyse DLA clusters. From this analogy, one thinks of the DLA cluster as a charged conductor, and finds the potential by solving Laplace's equation around the cluster. This means that if we know the growth probability at a given site, then we know the charge at that site in the electrostatic analogy.

A multipole expansion for the electric potential in 2 dimensions is

$$\phi(\mathbf{z}) = \sum_{l=1}^{\infty} \frac{1}{l} q_l \mathbf{z}^{-l} + q_0 \ln(z) \tag{5.12}$$

where the multipole moments are

$$q_l = \oint (z(s))^l \zeta(s) ds. \tag{5.13}$$

The integral is around the cluster boundary and $\zeta(s)$ is the charge density at point $(x(s), y(s))$ on the cluster. Since the charge density on the cluster is given by the growth probability at that point, we may sample the charge and estimate the multipole moments by using probe particles (see sec. 5.3.1).

The l^{th} multipole moment is proportional to the l^{th} power of the cluster radius, and we need to normalise the moments. The multipole power is chosen here to be defined by

$$P_l = \frac{|q_l|^2}{R_{eff}^{2l}} \tag{5.14}$$

where R_{eff} is the effective radius of the DLA cluster, defined by

$$\ln(R_{eff}) = \oint \ln(|z(s)|) \zeta(s) ds. \tag{5.15}$$

This normalisation is chosen because the multipole powers are related to the Laurent

coefficients, B_l, of the conformal map of the cluster [25] by

$$\begin{aligned} P_2 &= -B_{-1}/R_{eff} \\ P_3 &= -B_{-2}/R_{eff} \\ P_4 &= -\left(B_{-3}/R_{eff} - \tfrac{3}{2}B_{-1}^2/R_{eff}^2\right) \\ &\cdots \end{aligned} \tag{5.16}$$

We have assumed that the monopole moment is zero ($P_1 = 0$), which means choosing the centre of charge to be the centre of our co-ordinate system.

The multipole moments may be considered for clusters grown in dimensions other than 2. In 3 dimensions, the spherical harmonics are used in place of $\mathbf{z}^{-k}$, and the multipole moments are defined by ([46], chapter 4)

$$q_{l,m} = \int r^l Y_{l,m}(\theta,\phi)\zeta(\mathbf{x})d\mathbf{x}. \tag{5.17}$$

We chose the multipole power to be

$$P_{l+1} = \frac{\sum_{m=-l}^{l} |q_{l,m}|^2}{(2l+1)R_{eff}^{2l}} \tag{5.18}$$

where the effective radius is

$$\frac{1}{R_{eff}} = \int \frac{1}{r}\zeta(\mathbf{x})d\mathbf{x}. \tag{5.19}$$

Note that this is a slightly non-standard definition of the multipole powers, since we are choosing P_1 give the centre of charge. This is from a desire to simplify comparison between the multipole powers as we have defined in 2D and 3D.

5.4.4 Phase space plots

Consider 2 quantities, (F, G), both of which are governed by the same correction to scaling exponent. It is simple to show that if we grow a cluster, and plot F against G for various values of N, all the points will lie on a straight line (F and G are said to be linearly dependent). Thus we have a test to look for a single correction to scaling exponent.

Figure 5.13 shows such a plot for the dipole power, P_2, against the relative penetration depth, $\frac{\xi}{R_{dep}}$. It is unclear whether the data fall onto a straight line. The values for the dipole power appear to dip quickly, suggesting that the correction to

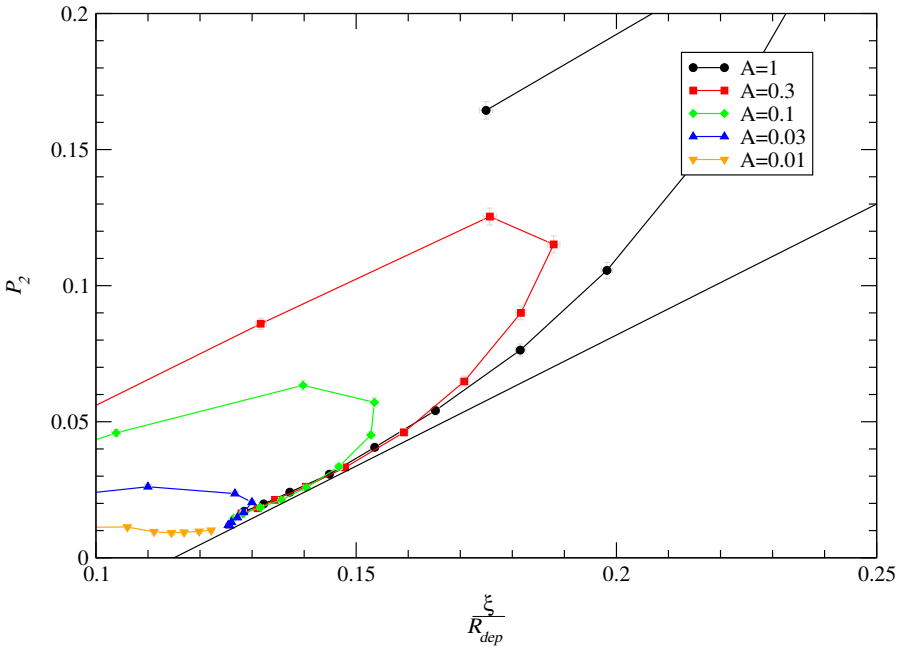

Figure 5.13: A phase space plot of the dipole power, P_2, against the relative penetration depth, $\frac{\xi}{R_{dep}}$. These quantities show a slight departure from a straight line for large cluster sizes. However, it is unclear whether this departure is significant.

scaling exponent for P_2 may be higher than for $\frac{\xi}{R_{dep}}$. However, it is difficult to know what is a significant departure from a straight line. With this plot it is not possible to extract the value for the correction to scaling exponent, ν.

5.4.5 Differential plots

To extract the correction to scaling exponents in an unbiased way, a "differential plot" is proposed, in which one plots a quantity against its logarithmic derivative.

Definition and why it's a good plot

Consider that we have some quantity, Q, which displays a single correction to scaling (eq. 5.10). If we differentiate this quantity with respect to the log of the cluster size, we find

$$\begin{aligned} \frac{dQ}{d\ln(N)} &= N\tfrac{d}{dN}(Q_\infty + CN^{-\nu}) \\ &= -\nu C N^{-\nu}. \end{aligned} \tag{5.20}$$

This leads to

$$\frac{dQ}{d\ln(N)} = -\nu(Q - Q_\infty). \tag{5.21}$$

If we plot $\frac{dQ}{d\ln(N)}$ against Q, we expect the data to fall onto a line of slope $-\nu$, which intercepts the x-axis at Q_∞, the asymptotic value.

The differential plots show, in an unbiased way, whether noise reduced DLA is the same as normal DLA. The hope is that noise reduced DLA will develop the asymptotic properties of DLA much more quickly than normal DLA. One may check whether noise reduced DLA possesses the same correction to scaling as normal DLA, as well as the same asymptotic properties. Unfortunately, the differentiation will amplify any noise present in the measurements since the difference between two adjacent measurements is expected to be small, but the error in the difference will be greater than the error in either of the measurements (assuming that the errors are uncorrelated). The differential is approximated by

$$\frac{dQ}{d\ln(N)} \simeq \frac{Q_{N_2} - Q_{N_1}}{\ln\left(\frac{N_2}{N_1}\right)} \tag{5.22}$$

and the error in the differential by

$$\sigma\left(\frac{dQ}{d\ln(N)}\right) \simeq \frac{\sqrt{\sigma^2(Q_{N_2}) + \sigma^2(Q_{N_1})}}{\ln\left(\frac{N_2}{N_1}\right)}. \tag{5.23}$$

Q_{N_1} is the measurement of Q for a cluster with N_1 particles. This has error $\sigma(Q_{N_1})$, which is measured as the standard deviation of Q_{N_1}, taken over r clusters, divided by $r - 1$.

5.5 Results

Figure 5.14 shows a differential plot for the relative penetration depth, $\frac{\xi}{R_{dep}}$. This gives a limiting value of $\left.\frac{\xi}{R_{dep}}\right|_\infty = 0.125 \pm 0.001$ and a correction to scaling exponent of $\nu = 0.35 \pm 0.07$. It should be noted that this provides a strong demonstration that our algorithm does produce noise reduced DLA. Not only is the penetration depth tending toward the same value for all choices of A, but it is approaching that value with the same correction to scaling exponent.

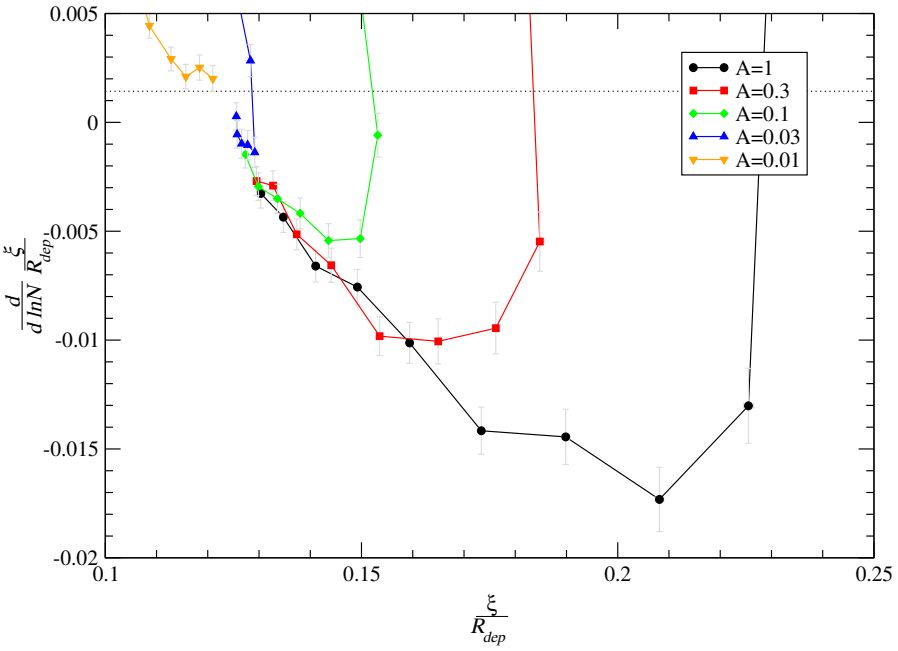

Figure 5.14: Differential plot for the relative penetration depth ($\frac{\xi}{R_{dep}}$). This shows an asymptotic value of $\left.\frac{\xi}{R_{dep}}\right|_{\infty} = 0.125 \pm 0.001$ and a correction to scaling exponent of $\nu = 0.35 \pm 0.07$.

Figure 5.15 shows a differential plot for the first 4 (non-trivial) multipole powers ($P_2 - P_5$, see eq. 5.14). They show similar values for the correction to scaling exponents.

The correction to scaling exponents for the multipole powers and the relative penetration depth are plotted in fig. 5.16. This data is consistent with the assertion that there is no quantity which exhibits a slower correction to scaling than the relative penetration depth, and the correction to scaling exponents for P_n, $n \leq 6$ are the same to within error. The data for $n > 6$ show influence from more rapidly decaying corrections, but it cannot be ruled out that these data also show the slowest correction to scaling. In particular, if the multipole powers are normalised differently (according to eq. 5.24) then the higher powers do suggest a much lower correction to scaling exponent.

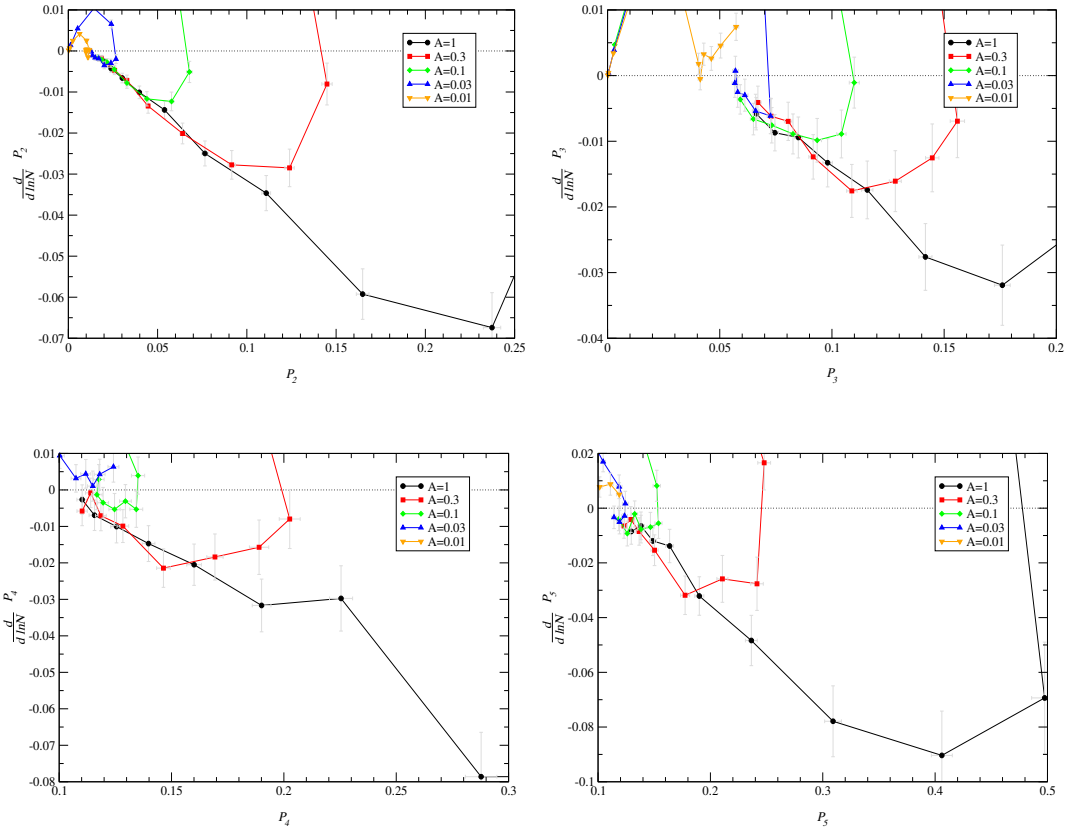

Figure 5.15: Differential plots for the first four (non-trivial) multipole moments $(P_2 - P_5)$. These show correction to scaling exponents as follows: $\nu(P_2) = 0.37 \pm 0.02$, $\nu(P_3) = 0.30 \pm 0.04$, $\nu(P_4) = 0.38 \pm 0.06$, and $\nu(P_5) = 0.35 \pm 0.03$. These gradients are estimated by eye, and the errors represent the maximum believable error.

5.5.1 Number of arms

In section 5.4.3, we chose the multipole powers to be normalised using R_{eff}. We now choose to normalise the multipole powers as follows

$$P_l' = \frac{|q_l|^2}{\left(\oint |z(s)|^l \zeta(s) ds\right)^2} \tag{5.24}$$

where the integral is around the boundary of the cluster. Using this definition we can measure the number of growing arms of the cluster. Ball [4] estimated that a DLA cluster would have 4.9 dominant growing arms, and measurements [51] have suggested this estimate to be accurate. With the multipole powers we have a direct way of testing this estimate, and in fig. 5.17 the asymptotic value of these multipole powers is plotted. The asymptotic values first reach their final value for the 4th multipole power, suggesting that DLA clusters have 4 arms. The data is expected to saturate, since the higher powers will be responding to the fine structure of the

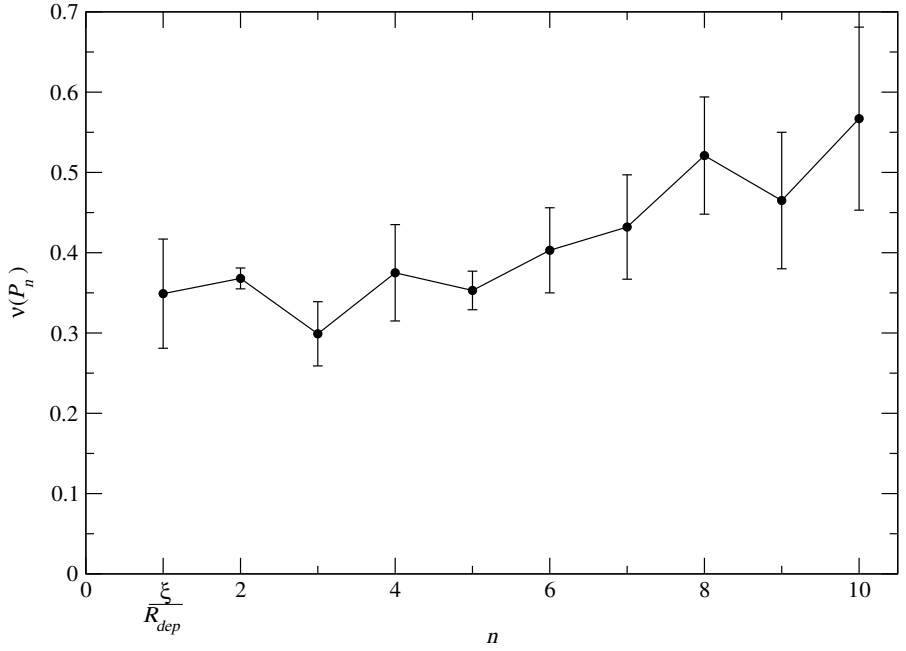

Figure 5.16: The correction to scaling exponents for the multipole powers $(P_2 - P_{10})$ and the relative penetration depth. The exponents for the higher multipole powers appear to be significantly larger than the exponent for either P_2 or $\frac{\xi}{R_{dep}}$.

cluster.

5.6 Conclusion

We have considered the development of DLA, and seen that noise reduction was crucial to the understanding of DLA on-lattice. A version of noise reduction off-lattice has been proposed, and its relation to the HL method recognised. This relation has been tested, and it was suggested that "fjord-filling" is catastrophic for the HL method.

The correction to scaling exponents of DLA have been found by use of differential plots. The high multipole powers show evidence of faster corrections to scaling than the slowest. Our measurements of the number of arms of a DLA cluster are consistent with previous measurements.

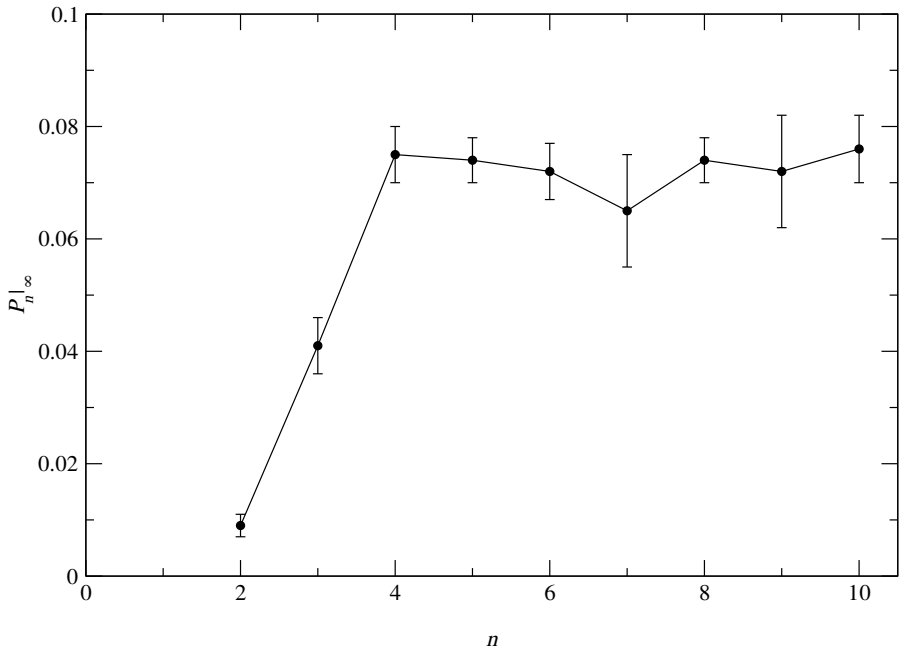

Figure 5.17: The asymptotic values for the multipole moments ($P_k^{'}$) when normalised by r^k. This suggests that a DLA cluster having 4 arms.

Chapter 6

Renormalisation and DLA in 3D

I do not feel obliged to believe that the same God
who has endowed us with sense, reason, and intellect
has intended us to forgo their use. – Galileo Galilei

In this chapter, renormalisation approaches to the understanding of DLA are described, and we consider the properties of DLA clusters grown in 3 dimensions. We identify a fixed point of the renormalisation scheme, which corresponds to the asymptotic limit of large clusters. 3D DLA clusters are little understood, and it is hoped that results in this field can illustrate the distinctions between 2D and 3D DLA.

6.1 Introduction to renormalisation

We have considered a fractal to be defined by self-similarity (sec. 5.1.3). This idea naturally leads to the concept of renormalisation. If we have a shape that is composed of a number of copies of itself, consider formally changing the length scale at which we are looking; carefully renormalising all the appropriate quantities so that the change of length scale does not generate a change in structure.

This section is composed of a review of previous renormalisation schemes. The concept of a renormalised noise reduction is applied to the data for noise reduced DLA, and checked against the fixed point which is immediately apparent. This poses some difficulties and an effective noise reduction is introduced to explain this problem.

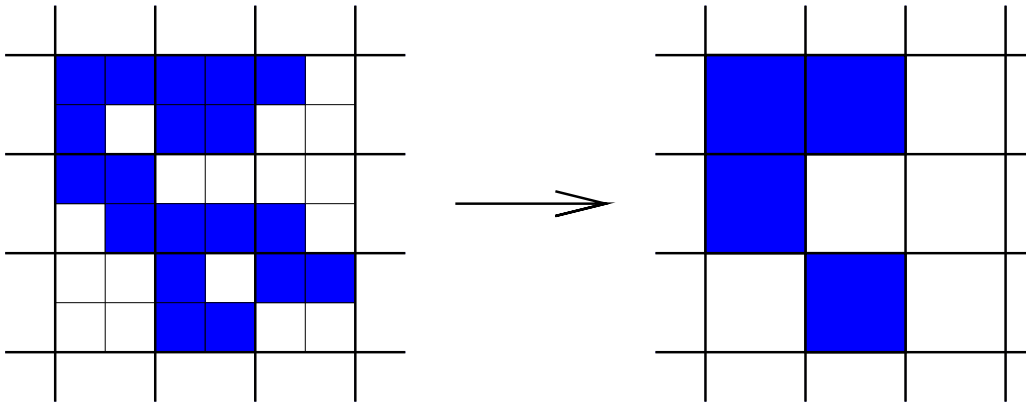

Figure 6.1: A simple renormalisation scheme. Each set of 4 boxes are replaced by a "super box". The super box is occupied only if 3 or more of the smaller boxes are as well.

6.1.1 Renormalisation in general

In the renormalisation approach, we view a system at different scales. In the context of DLA, we consider replacing a portion of the cluster with a single particle, whose properties appropriately match the portion replaced.

Consider a simple renormalisation scheme for an on-lattice DLA cluster (see fig. 6.1). We imagine replacing each group of 4 sites by a super site. This super site will be occupied if at least 3 of the individual sites are occupied. This will create a coarse-grained cluster, whose properties may differ from the original cluster. If it happens that the properties of the cluster are not altered by the renormalisation, we are said to be at a fixed point of the renormalisation scheme.

Being at a fixed point of the renormalisation implies the cluster being scale-invariant. Let us consider some property of the cluster, q. At a fixed point q takes the value

$$q^* = R(q^*) \tag{6.1}$$

where R is a function which describes how q changes under the renormalisation. If we are near this fixed point, then

$$q^* + \delta q' = R(q^* + \delta q) \tag{6.2}$$

where the system takes the value $q^* + \delta q$ before the renormalisation and $q^* + \delta q'$ after. Taylor expansion of R gives

$$q^* + \delta q' = R(q^*) + \delta q R'(q^*) + \ldots \tag{6.3}$$

This means that the deviation of the system from the fixed point transforms as

$$\delta q \to C\delta q \tag{6.4}$$

where $C = R'(q^*)$ is a constant. Since this renormalisation involves a change of scale, it also reduces the size of the cluster by a constant factor

$$N \to DN. \tag{6.5}$$

This leads to

$$\delta q = AN^{\frac{\ln(C)}{\ln(D)}} \tag{6.6}$$

Hence q shows a correction to scaling. In general, we consider R to be a function of the many parameters that describe the system. We may write the change in the system under renormalisation in the following form

$$\mathbf{q}' = \mathbf{R}\mathbf{q} \tag{6.7}$$

where $\mathbf{q} = (q_1, q_2, \ldots, q_k)$ are the parameters describing the system. We may linearise around the fixed point, the eigenvalues of the matrix $\mathbf{R}$ giving the correction to scaling exponents. One notes that there may be no more correction to scaling exponents than parameters in the renormalisation.

6.1.2 Previous work

Barker and Ball scheme

Barker & Ball [7] proposed a renormalisation scheme for DLA (on-lattice) where the noise reduction and anisotropy are the key parameters. This led them to conlcude that the fixed point noise reduction for DLA is $\epsilon = \frac{1}{m} \sim 0.01$.

The renormalisation is based on a DLA algorithm [115, 113, 3] which is an extension of the original Witten & Sander model. A diffusing particle performs a random walk until it comes close to the cluster. If it lands on a site neighbouring the cluster, it sticks. If it lands on a site which is a next-nearest neighbour to the cluster, it sticks with probability p. If it does not stick, then it continues to diffuse until it becomes stuck. This introduces an anisotropy, which competes with the lattice anisotropy. In the limit $p = 0$, we get growth along the axes, and for $p = 1$, we see growth along the diagonals [5].

They grew clusters to a given (extremal) radius, and measured the properties these clusters. The two properties of interest are considered to be the noise reduction and anisotropy of the cluster grown. These measured properties can then be used to grow another set of clusters, with the measured values of the anisotropy and noise reduction used as input parameters when growing the new clusters. The noise reduction and anisotropy were measured using

$$\epsilon_L = \frac{1}{< N_w >} + \frac{< N_w^2 > - < N_w >^2}{< N_w >^2} \tag{6.8}$$

$$p_L = \langle min(\cot(\theta), \tan(\theta)) \rangle \tag{6.9}$$

where N_w is the number of diffusing particles required to grow the cluster to a given radius, the average is taken over an ensemble of clusters, and θ is the angle (with respect to the origin) at which the cluster first reaches the given radius. Eq. 6.8 follows from the time taken to grow a given (lattice) cell in a cluster being governed by a Poisson process. It is assumed that the distribution of the time taken to grow a cluster possesses the same structure, and that this equation may therefore be applied to entire clusters. Barker & Ball performed simulations of clusters in a quadrant, growing to radii of 4.5, 7.5, and 11.5 lattice units. They concluded that the isotropic fixed point for DLA is an unstable fixed point, given by

$$\epsilon_\infty = 0.0064 \pm 0.0004 \tag{6.10}$$

$$p_\infty = 0.464 \pm 0.002. \tag{6.11}$$

FST scheme

An alternative approach to the renormalisation of DLA was proposed by Cafiero et al [19]. Rather than consider the normal DLA algorithm, they considered the bond model of DLA [81], without the possibility of growth on the diagonal. They argued that this case was the correct one to consider, since under renormalisation the ordinary DLA algorithm would transform into this case. The DLA clusters considered were in a cylindrical geometry, and the properties being measured only in the "dead zone" of the cluster (see sec. 5.1.7). This, they argued, is the correct geometry, as the active zone of DLA only serves to complicate the situation.

The scheme considers the correlation between pairs of adjacent cells, each of size $L * L$. These correlations are then used to calculate the noise reduction via eq.

6.8. This was performed analytically for $L = 1$ [19]. They found an attractive fixed point noise reduction of $m \simeq 2.4$. This result was essentially confirmed by a Monte Carlo simulation for L up to 512 [89]. They claim that this explains the fractal nature of DLA, since screening is clearly present in DLA with low noise reduction.

The number of correction to scaling exponents

In Cafiero et al. [19], they considered a renormalisation scheme based on only one parameter. Similarly, if we wish to apply the Barker & Ball [7] off-lattice, then we will be considering only the noise reduction, and we can only relate this parameter to one correction to scaling exponent. If there is more than one correction to scaling exponent describing the growth of DLA around the fixed point, then the renormalisation scheme needs to be extended to account for this. This may indicate that the response of DLA to anisotropy needs to be (re-)introduced.

6.2 Measurements of the fixed point noise reduction

6.2.1 The effect of noise reduction

It is clear from differential plots such as figure 5.14 that the noise reduction "controls" the slowest correction to scaling. For low values of A this correction to scaling is strongly reduced, and we may write the behaviour of this correction as follows

$$Q = Q_\infty + C(A)N^{-\nu}. \tag{6.12}$$

The other corrections to scaling need not depend on A, but it is quite evident that the slowest correction to scaling is strongly affected by A.

If for some value of A we have $C(A_f) = 0$, then clusters grown at this value of the noise reduction will approach the fixed point most quickly. These clusters will "short-circuit" much of the evolution of a DLA cluster. If we grow clusters with a value of A equal to the fixed point value of ϵ^*, then we expect that this will also "short-circuit" much of the evolution of a DLA cluster. Hence, we associate the value of A which arrives fastest at the fixed point with the fixed point noise reduction.

6.2.2 Fixed point by direct observation

Since differential plots allow us to easily measure the slowest correction to scaling exponent, we may also identify A_f. This is the value of A which does not show the slowest correction to scaling, and hence approaches the fixed point from a different angle on the plot. Figure 5.14 for the relative penetration depth indicates that $0.01 < A_f < 0.03$. Figure 5.15 for P_2 suggests that $A_f \simeq 0.01$ may be more appropriate.

Although the data presented here and of Ellak Somfai [103] are indistinguishable for most measurements, they disagree slightly for the value of A_f. Somfai's data suggests a slightly lower value for A_f than the data presented here. It is clear that the difference between the two sets of data is small, but that difference is noticeable when the value of A_f is considered. Hence, $A_f \simeq 0.01$ will be used in the following analysis.

6.2.3 Fixed point from renormalisation

We consider measurements of the fixed point noise reduction using the Barker & Ball [7] formula, eq. 6.8. In the following A will be used to describe the noise reduction used in the microscopic growth rules (i.e. A_f) and ϵ will denote the noise reduction as measured from cluster properties. Since all the measurements presented here have been taken at a fixed value of N, eq. 6.8 needs to be recast into an appropriate form. The fluctuations in R and N are illustrated in figure 6.2. Since the cluster scales with dimension D, we must have

$$\ln\left(1+\frac{\delta R}{R}\right) = -\frac{1}{D}\ln\left(1-\frac{\delta N}{N}\right). \tag{6.13}$$

Taylor expansion to first order gives

$$\frac{\delta N}{N} = D\frac{\delta R}{R}. \tag{6.14}$$

So we expect that the fixed point noise reduction will be

$$\epsilon^* = D^2\left(\frac{\delta R_{ext}}{R_{ext}}\right)^2, \tag{6.15}$$

taking $\frac{1}{<N>}$ to be small.

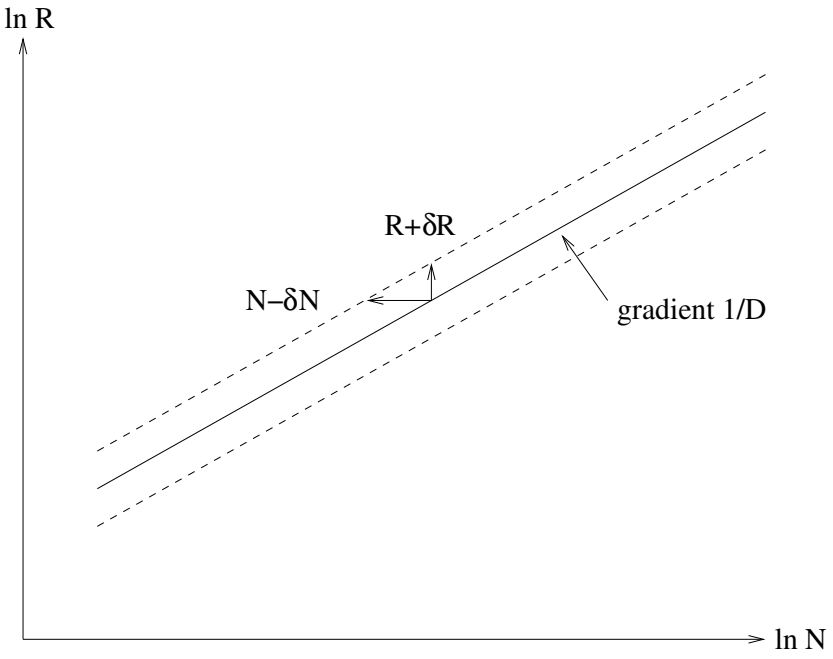

Figure 6.2: Relation between the fluctuations in N and R. Since the cluster scales with dimension D, we must have $\ln(R+\delta R) - \ln(R) = (1/D)(\ln(N) - \ln(N-\delta N))$.

Figure 6.3 shows a differential plot for the relative variability of R_{ext}. From this the asymptotic value is measured to be

$$\left.\frac{\delta R_{ext}}{R_{ext}}\right|_{N\to\infty} = 0.034 \pm 0.005. \tag{6.16}$$

This leads to an estimate of the fixed point noise reduction of

$$\epsilon^* = 0.0034 \pm 0.0010. \tag{6.17}$$

This estimate is similar to the value found by Barker & Ball [7], and in clear disagreement with the value of Cafiero et al. [19]. We are therefore forced to conclude that either the Cafiero analysis is in error, or that radial DLA is intrinsically different from cylindrical DLA. Since we have already seen that the penetration depth is a significant fraction of the radius, it is difficult to make equivalent measurements of the two systems. For example, exactly how one may make an equivalent measurement of ϵ in cylindrical geometry is not clear. Further, it is not known whether clusters grown in these two geometries have identical properties [32], and more simulations are needed in this area.

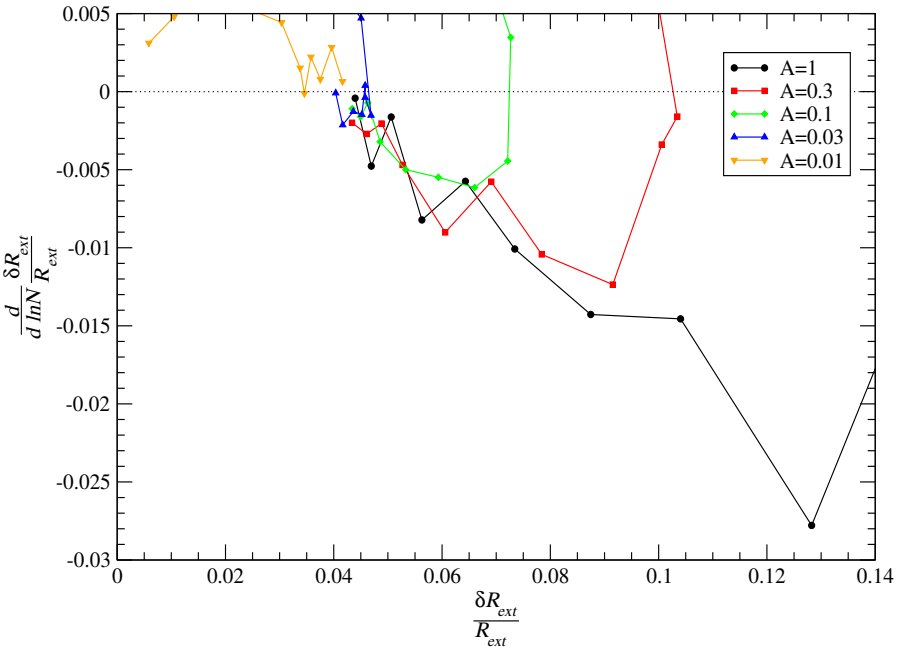

Figure 6.3: A differential plot of the relative variability of the extremal cluster radius (measured with respect to the seed). The asymptotic value is $\frac{\delta R_{ext}}{R_{ext}}\Big|_{N\to\infty} = 0.034 \pm 0.005$. This leads to an estimate of the fixed point noise reduction of $\epsilon^* = 0.0034 \pm 0.0010$.

6.3 Effective noise reduction

Given the discrepancy between ϵ^* and A_f, we consider the number particles required to grow the cluster to a given value of the average radius of deposition. Figure 6.4 shows $R_{dep}(N)$ against $A * N$, and we might expect that this quantity tends toward the same curve for all values of A. This is not that case, even for large clusters. $R_{dep}(N)$ is smaller than expected for small values of A. We conclude that it requires more than $\frac{1}{A}$ particles to grow the equivalent of 1 particle at $A = 1$, leading to the concept of an effective noise reduction for each value of A.

This problem was apparent in on-lattice DLA [31], when it was noticed that the clusters grown exhibited compact growth for a period before branching occurred. One may note that this effect occurs for all branches and sub-branches and is illustrated in figure 6.5. Until two branches are separated by more than 1 particle diameter, no diffusing particle can penetrate between them. Since the diffusive field forces them to grow in nearly the same direction, they persist in being close for

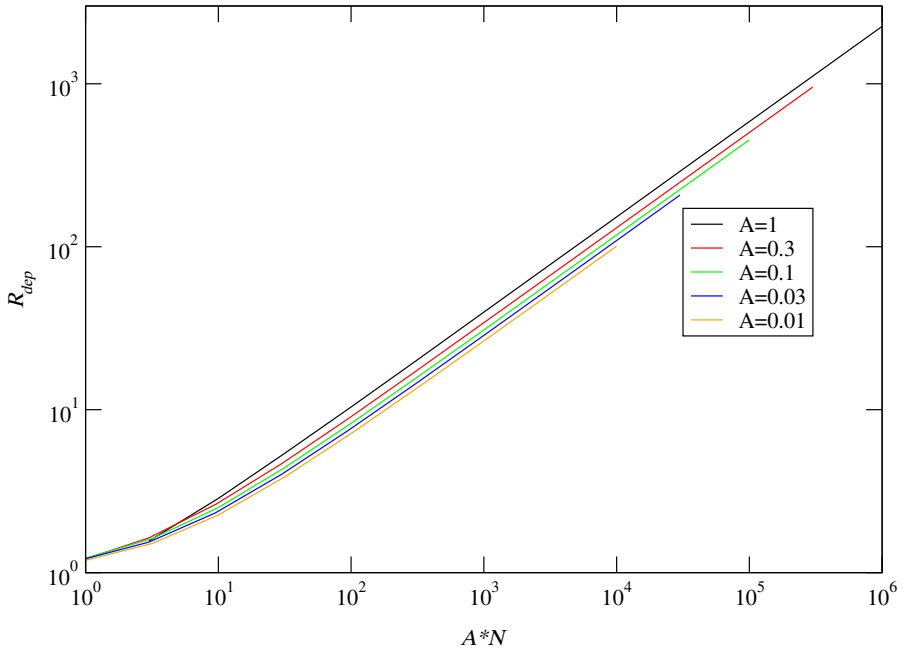

Figure 6.4: R_{dep} against $A * N$ for each value of A. The clusters grown with small values of A are more compact than expected showing that it requires more than $\frac{1}{A}$ particles to grow the equivalent of 1 particle at $A = 1$.

a long time. This effect persists longer for smaller values of A, because a single particle depositing on the side of a tip will affect the growth less and the branches will become distinct more slowly.

6.3.1 2 point correlation function

We use the 2 point correlation function (eq. 5.5) as a way of measuring the effective noise reduction, ϵ^{eff}, for a given value of A, since C_2 is independent of the location of the seed. $C_2(r)\epsilon^{eff}$ is plotted against r for 5 clusters in figure 6.6, ϵ^{eff} being chosen so that a data collapse is seen for small r. The values of the effective noise reduction are found to be

$$\begin{aligned} A = 1 \quad & \epsilon^{eff} = 1 \\ A = 0.3 \quad & \epsilon^{eff} = 0.24 \\ A = 0.1 \quad & \epsilon^{eff} = 0.068 \\ A = 0.03 \quad & \epsilon^{eff} = 0.018 \\ A = 0.01 \quad & \epsilon^{eff} = 0.0054 \end{aligned} \tag{6.18}$$

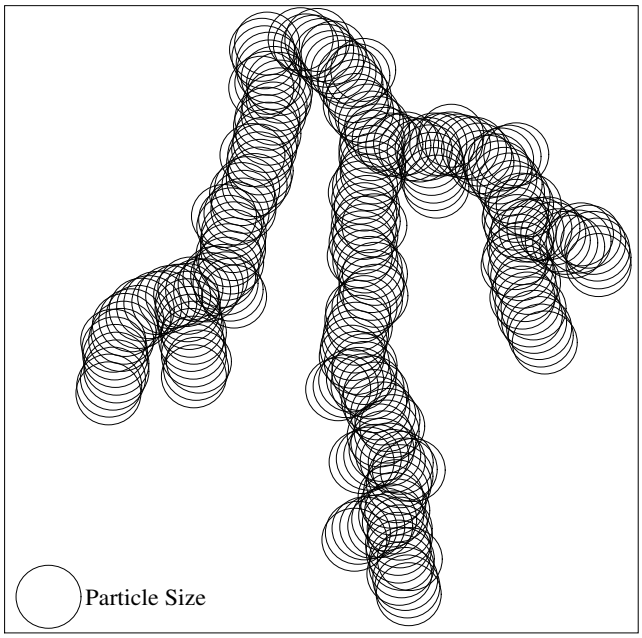

Figure 6.5: Compact growth of two neighbouring arms. The two arms are closer than one particle diameter for much of the development of the arm. In this case, they are effectively growing a compact structure. Since the diffusive field forces them to grow in nearly the same direction, they persist in being close for a long time. It is only when a number of particles deposit on the side of the tip that they grow apart.

The fixed point noise reduction is therefore estimated to be

$$\epsilon_f^{eff} \simeq 0.0054 \tag{6.19}$$

which is much closer to the measured value of ϵ^*.

6.3.2 Branch distribution

As another measure of the effective noise reduction, a new quantity is defined: the branch distribution, which is the number of branches, $N_{branch}(L)$, which are longer than a length L. The length of a branch is measured as the Euclidean distance between the branch point at which it was created (or the seed) and its tip. The smallest branches on the cluster are sequentially pruned before the lengths of the longer branches are considered (see fig. 6.7).

Figure 6.8 shows the branch distribution divided by the effective noise reduction $\frac{N_{branch}(L_{eff})}{\epsilon^{eff}}$ for five clusters against an effective length, L_{eff}. The graphs

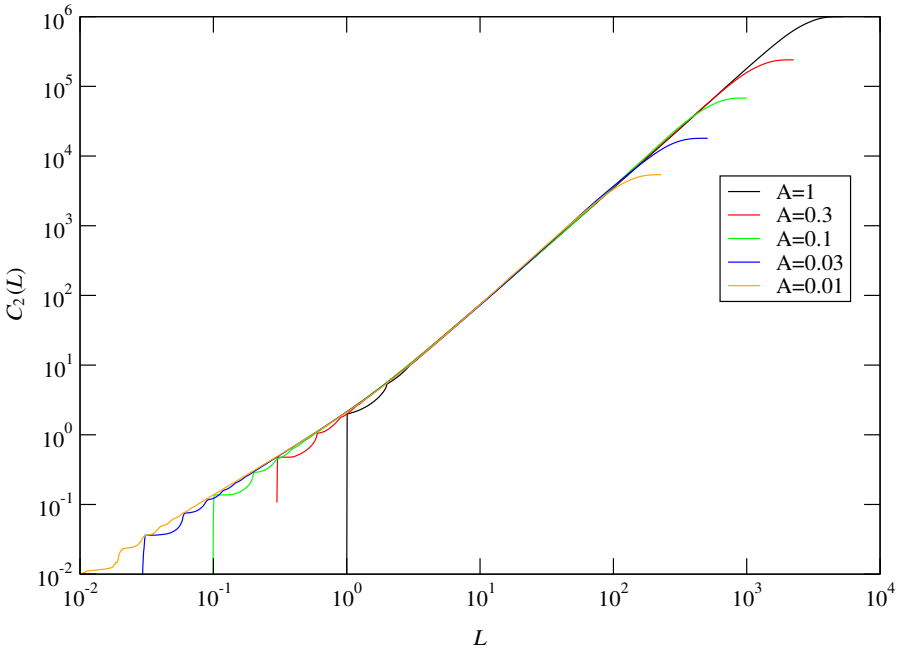

Figure 6.6: The 2 point correlation function multiplied by an effective noise reduction, ϵ^{eff} for clusters grown in 2D. A data collapse is achieved for small length scales by choosing $\epsilon^{eff} < A$ for small values of A, showing that noise reduced DLA clusters are more compact than one might initially expect.

for $A < 1$ have been shifted so that their values for small L_{eff} fall onto the same curve. The shift factors are

$$\begin{array}{lll} A = 1 & \epsilon^{eff} = 1 & L_{eff} = L \\ A = 0.3 & \epsilon^{eff} = 0.21 & L_{eff} = 0.8L \\ A = 0.1 & \epsilon^{eff} = 0.056 & L_{eff} = 0.65L \\ A = 0.03 & \epsilon^{eff} = 0.013 & L_{eff} = 0.5L \\ A = 0.01 & \epsilon^{eff} = 0.0033 & L_{eff} = 0.4L \end{array} \tag{6.20}$$

where L is the real length of a given branch. This leads us to estimate the fixed point noise reduction, ϵ_f^{eff}, to be

$$\epsilon_f^{eff} \simeq 0.0033 \tag{6.21}$$

This value agrees well with the measured value of ϵ^*. It is, however, different from the value given by the two point correlation function. Both methods for evaluating the effective noise reduction give values for ϵ_f^{eff} which are close to the value of ϵ^*.

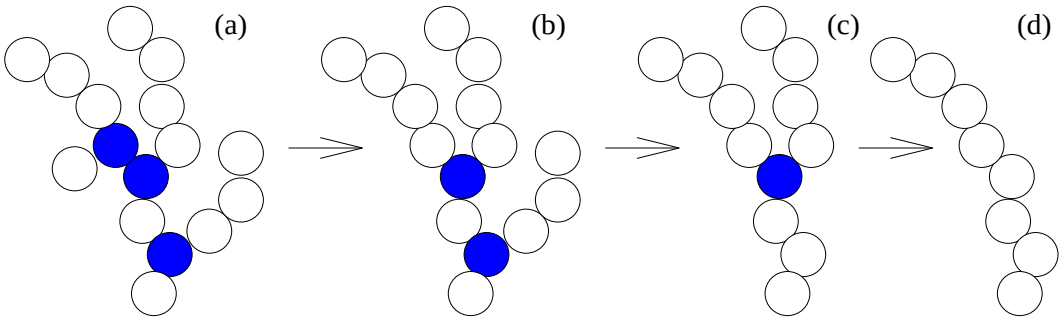

Figure 6.7: How the length of branches are calculated. The length of each branch is given by the Euclidean distance between the tip of the branch and its branch point (shaded). The shortest branches are trimmed first, so that there is an iterative reassessment of the length of each branch. In the original picture (a) the length of the shortest branch is recorded, and branch pruned (b). The length of the branch that is now shortest is recorded, and this deleted (c). This is repeated until a single branch is left which is much longer than any branch in the original picture (a).

6.4 DLA clusters grown in 3D

Much of the work on DLA has been restricted to 2 dimensions as this case has proved sufficiently difficult. Off-lattice noise reduction has been used to grow DLA clusters in 3 dimensions, and a number of interesting results have been found. Previously, noise reduction off-lattice has been restricted to 2 dimensions, since the HL method relies on a conformal map. The method described above is applicable in all dimensions.

6.4.1 Previous work on DLA in 3D

Most of the work on DLA has been in 2 dimensions. Meakin has been principle in extending work in DLA to higher dimensions, growing DLA clusters in dimensions up to 8 [111]. Much of that work has focused on estimating the fractal dimension of DLA clusters, and the scaling of the penetration depth [74, 73, 98]. Progress has also been made in measuring the multifractal spectrum ($f(\alpha)$ see [33], page 76) of DLA in 3D [101, 100, 36]. However, since Jensen et al. [49] have recently claimed that all previous attempts to measure $f(\alpha)$ in 2D have been flawed, it is expected that such attempts in 3D are also lacking. Other work has also examined DLA on a cubic lattice in the limit of zero noise [9], and the extension of the fixed scale transformation to 3 dimensions [112].

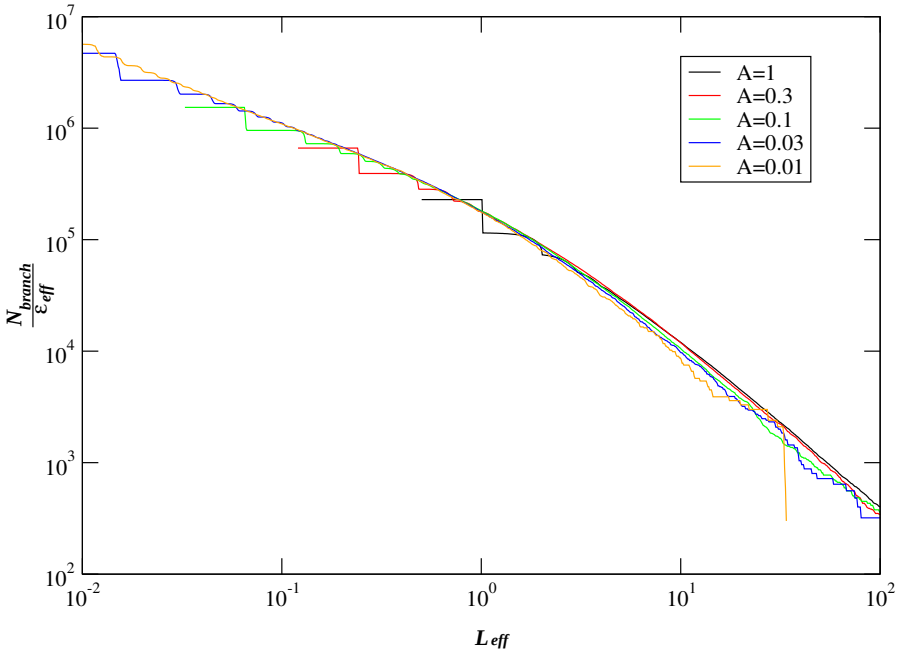

Figure 6.8: The branch distribution divided by an effective noise reduction, ϵ^{eff}, plotted against an effective length, L_{eff}. ϵ^{eff} and L_{eff} are chosen to give a good data collapse for small effective lengths. To achieve this it is found that $\epsilon^{eff} < A$ and $L_{eff} < L$ for small values of A.

6.4.2 Fractal dimension

The fractal dimension may be estimated by the local slope method outlined in section 5.2.3. The results are shown in figure 6.10. The dimensions are converging to a value of $D = 2.50 \pm 0.01$. The data for $A = 0.03$ and $A = 0.01$ are poorly converged, and the results could be made significantly more accurate with data for larger N.

6.5 Correction to scaling exponents in 3D

We now consider the correction to scaling exponents for DLA in 3 dimensions. In 2D, it has been proposed that DLA should exhibit a single, slowest correction to scaling [104]. Since we have only one control parameter for DLA, one might hope that this is the case. However, our results for DLA in 2D showed that the higher multipole powers are influenced by faster decaying corrections. Is this situation is mimicked in 3D, or do 3D DLA clusters show a different behaviour?

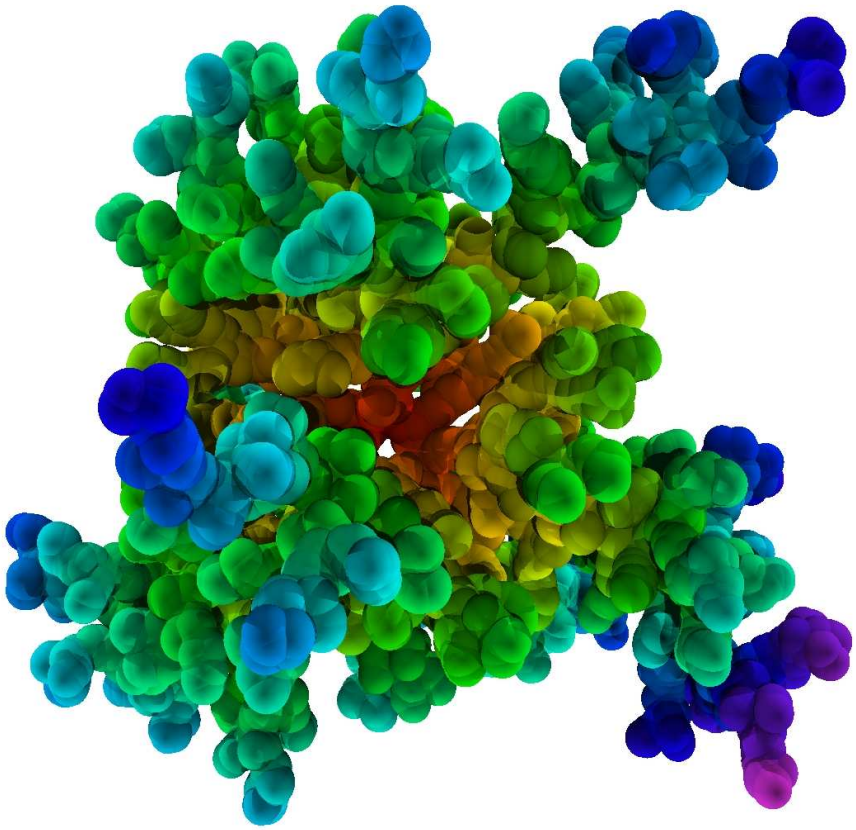

Figure 6.9: A 3D DLA cluster with $N = 10^4$ and $A = 0.1$, where the different shading indicates a different time of deposition on the cluster.

To extract the correction to scaling exponents for 3D DLA, differential plots for the relative penetration depth (fig. 6.11), multipole powers (P_2 - P_{10}, see eq. 5.18) (P_2 - P_5 are shown in fig. 6.12) and the relative variability of extremal cluster radius (fig. 6.13) are used. Even though conformal maps do not exist in 3 dimensions, the multipole powers are still a useful characterisation.

Figure 6.14 shows the measured values of the correction to scaling exponents for the above quantities. The exponent for P_2 is significantly different from the exponents for $\frac{\xi}{R_{dep}}$ and P_3. There is no quantity which shows a slower correction to scaling than $\frac{\xi}{R_{dep}}$, which suggests that the result found by Somfai et al. [104] (see sec. 5.4.2) in some sense also applies in 3D clusters. As in 2D, P_3 clearly shows the slowest correction to scaling of all the multipole powers. The continuous drift of the exponents of P_n to higher values as n increases seen in 2D does not appear in 3D, similar to the multipole powers, when normalised by $< r^n >$.

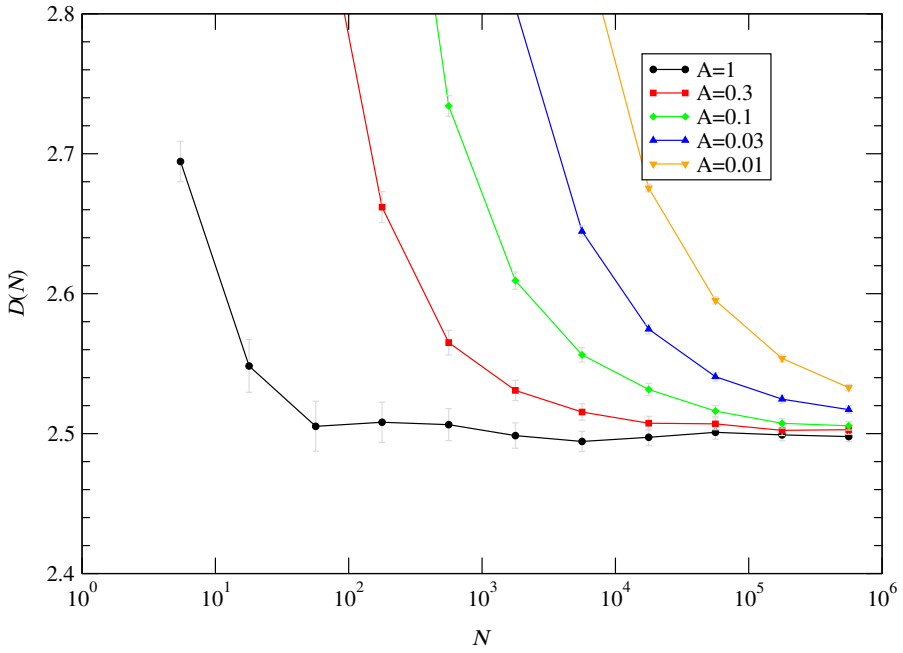

Figure 6.10: The measurement of the fractal dimension of clusters grown in 3D, by taking the local slope of R_{dep}. The dimension appears to be converging to a dimension of $D = 2.50 \pm 0.01$.

As with the results in 2D, the data are least precise for the higher multipole moments, as these are more sensitive to the fine structure of the cluster. The correction to scaling exponents appear to be smaller in 3D than in 2D. This confirms that 3D clusters are slower to mature than 2D clusters.

In spite of the differences, there is one aspect of 3D clusters that is strikingly similar to 2D clusters - the relative penetration depth. The asymptotic values of the relative penetration depth of 2D and 3D clusters are equal, to within measurement error.

6.6 Fixed point noise reduction in 3D

6.6.1 From direct observation

From differential plots (figs. 6.11 and 6.12) we can identify a fixed point noise reduction of the growth. The plots for $P_2 - P_4$ suggest that $A_f^{3D} < 0.01$ and plots of $\frac{\xi}{R_{dep}}, P_5$ are unclear as to the value of the fixed point noise reduction. Hence, one

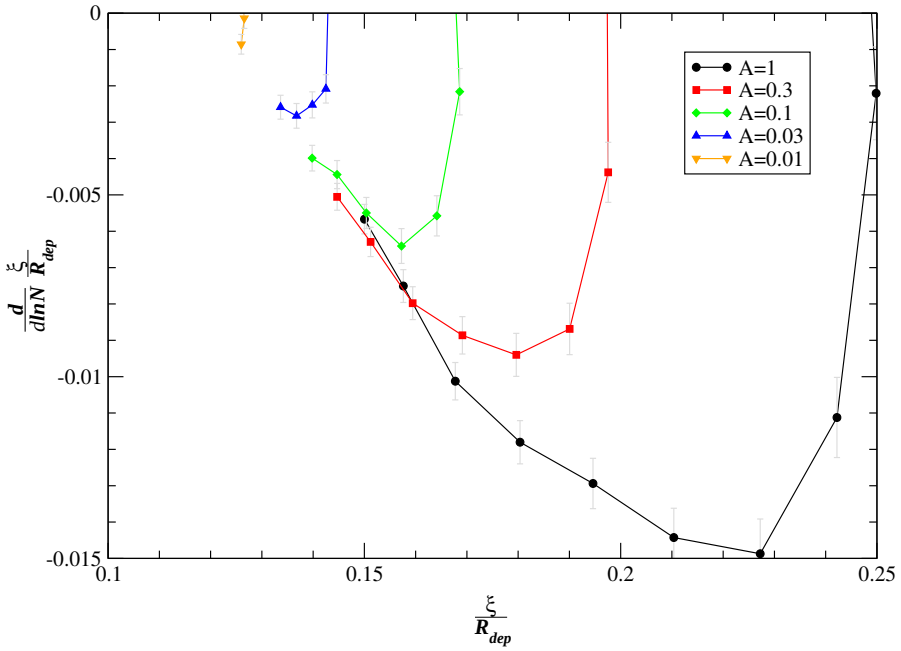

Figure 6.11: Differential plot of $\frac{\xi}{R_{dep}}$. The asymptotic value is $\frac{\xi}{R_{dep}}\Big|_{\infty} = 0.122 \pm 0.002$ and a correction to scaling exponent of $\nu = 0.22 \pm 0.03$. Note that clusters grown in 3 dimensions appear to take longer to mature.

estimates that

$$A_f^{3D} \leq 0.01. \tag{6.22}$$

6.6.2 From renormalisation

From figure 6.13, the asymptotic value of the relative variability of extremal cluster radius is $\frac{\delta R_{ext}}{R_{ext}}\Big|_{\infty} = 0.032 \pm 0.004$. This leads to

$$\epsilon_{3D}^{*} = 0.0064 \pm 0.0016. \tag{6.23}$$

This is close to the estimated value of A_f^{3D}. However, we need to take into account that each value of A corresponds to a value of ϵ^{eff}, the effective noise reduction.

6.6.3 C2 and N_{branch}

Graphs for the two point correlation function and branch distribution are shown in figures 6.15 and 6.16. These show shift factors very similar to those for 2D clusters,

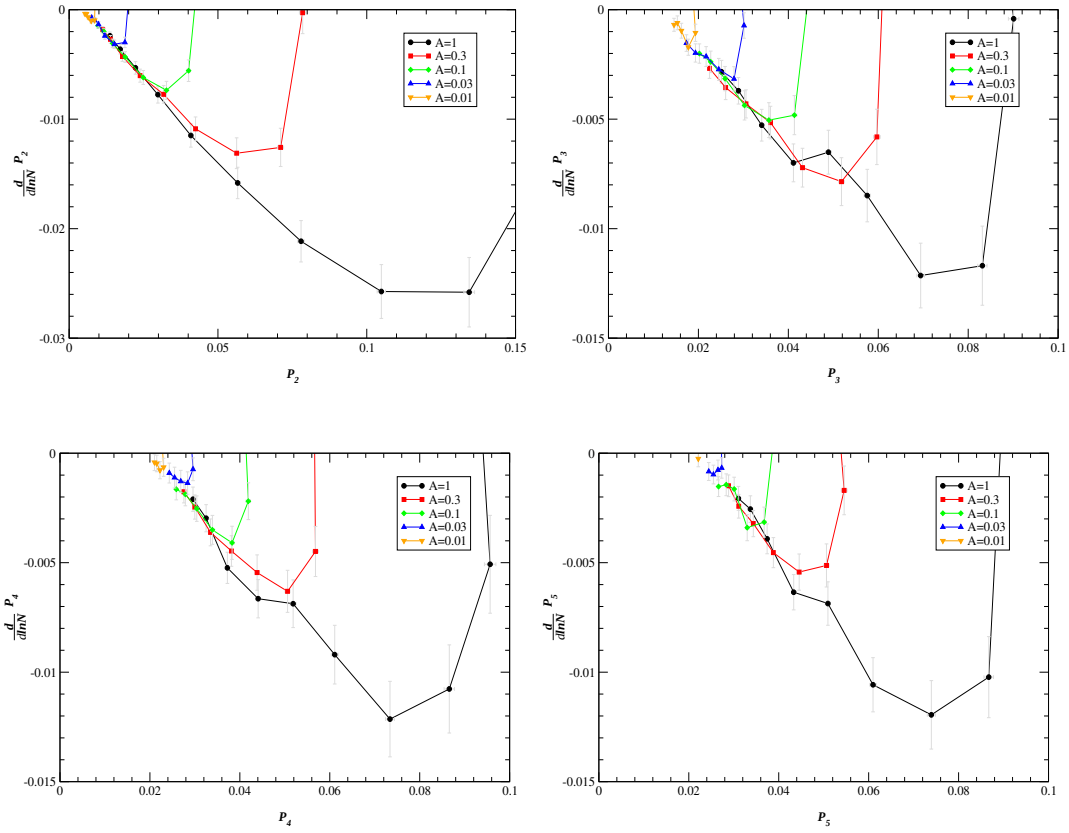

Figure 6.12: Differential plots for the first four (non-trivial) multipole moments (P_2–P_5). These show correction to scaling exponents as follows: $\nu(P_2) = 0.32 \pm 0.02$, $\nu(P_3) = 0.24 \pm 0.03$, $\nu(P_4) = 0.26 \pm 0.06$, and $\nu(P_5) = 0.29 \pm 0.05$. These gradients are estimated by eye, and the errors represent the maximum believable error.

and they are

Shift factors from C_2

$$\begin{aligned} A = 1 \quad & \epsilon^{eff} = 1 \\ A = 0.3 \quad & \epsilon^{eff} = 0.19 \\ A = 0.1 \quad & \epsilon^{eff} = 0.05 \\ A = 0.03 \quad & \epsilon^{eff} = 0.012 \\ A = 0.01 \quad & \epsilon^{eff} = 0.0034 \end{aligned} \tag{6.24}$$

Shift factors from N_{branch}

$$\begin{aligned} A = 1 \quad & \epsilon^{eff} = 1 \quad & L_{eff} = 1L \\ A = 0.3 \quad & \epsilon^{eff} = 0.21 \quad & L_{eff} = 0.9L \\ A = 0.1 \quad & \epsilon^{eff} = 0.056 \quad & L_{eff} = 0.8L \\ A = 0.03 \quad & \epsilon^{eff} = 0.013 \quad & L_{eff} = 0.7L \\ A = 0.01 \quad & \epsilon^{eff} = 0.0035 \quad & L_{eff} = 0.6L. \end{aligned} \tag{6.25}$$

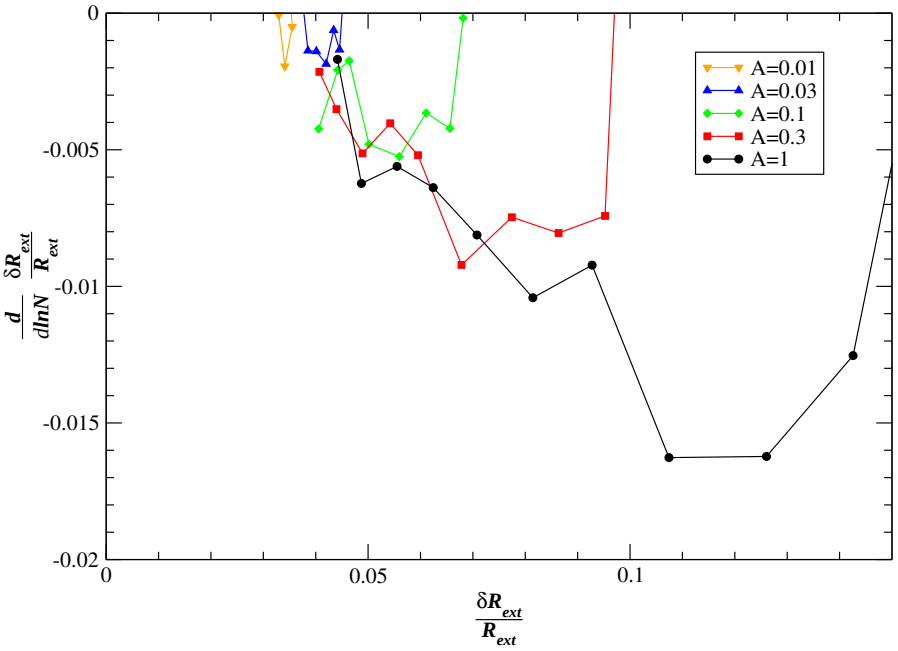

Figure 6.13: Differential plot of $\frac{\delta R_{ext}}{R_{ext}}$. The asymptotic value is $\frac{\delta R_{ext}}{R_{ext}}\Big|_{\infty} = 0.032 \pm 0.004$ which leads to an estimate of the fixed point noise reduction of $\epsilon^*_{3D} = 0.0064 \pm 0.0016$.

Hence one concludes that

$$\epsilon_f^{3D} \leq 0.0035. \tag{6.26}$$

Whereas the shift factors in 2D aid agreement between ϵ_f and ϵ^*, we find that in 3D they push the values further apart.

6.6.4 Are we worried?

We know that the identification of $\epsilon^* = D^2 \left(\frac{\delta R}{R}\right)^2$ is somewhat approximate. This identification gives a result which is within a factor of 2 of ϵ_f^{3D}, and this may be considered good agreement. Further, if we believe that there is no single correction to scaling exponent for DLA, then we have to reason to expect that a renormalisation scheme which uses a single parameter may be inaccurate.

6.6.5 Relative variability of the radius

So far, we have only considered the relative variability of the extremal cluster radius. Since the renormalisation scheme is based on the variability of the time it takes

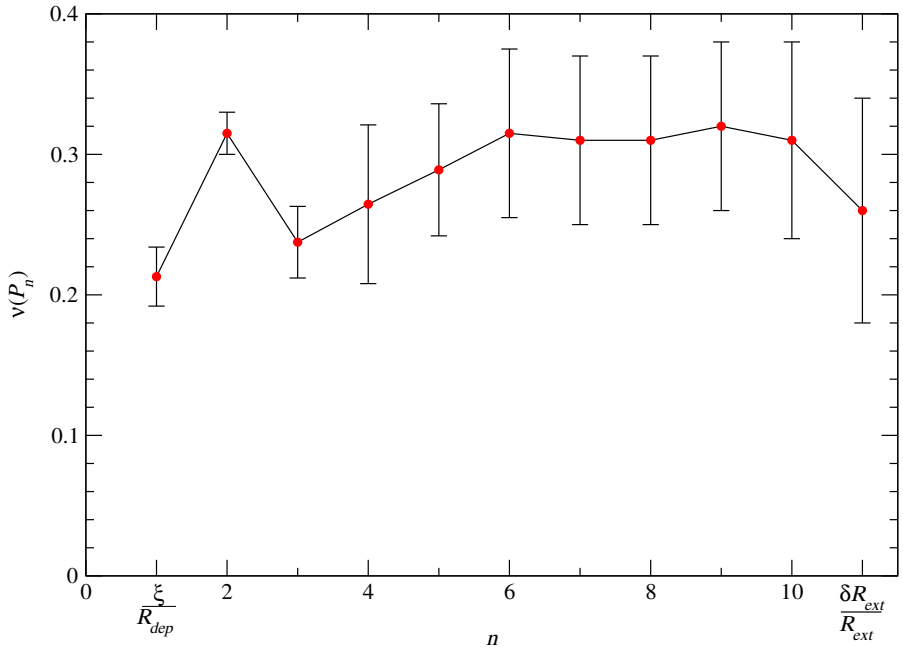

Figure 6.14: The correction to scaling exponents measured for 3D DLA clusters. The exponent for P_2 is significantly different from the exponents measured for $\frac{\xi}{R_{dep}}$ and P_3. The exponents for the higher multipole powers do not show the drift measured for the multipole powers in 2D.

the cluster to grow to a given radius, the extremal cluster radius is the natural quantity to use. However, in figure 6.17 the relative variability of R_{dep} is plotted. In 2D the asymptotic value is $\left.\frac{\delta R^{2D}_{dep}}{R^{2D}_{dep}}\right|_{\infty} = 0.014 \pm 0.001$, but in 3D it is $\left.\frac{\delta R^{3D}_{dep}}{R^{3D}_{dep}}\right|_{\infty} = 0.0067 \pm 0.0005$.

The difference in the relative variability of R_{dep} between 2D and 3D is surprising, since $\left.\frac{\delta R^{2D}_{ext}}{R^{2D}_{ext}}\right|_{\infty} \simeq \left.\frac{\delta R^{3D}_{ext}}{R^{3D}_{ext}}\right|_{\infty} \simeq 0.033$. So, we have a picture that for 3D clusters $\frac{\delta R_{ext}}{R_{ext}}$ is dominated by a few, thin branches that extend far from the cluster. Hence, we may wish to chose a slightly different radius, to ensure that a significant fraction of the cluster has reached that radius.

6.7 Number of arms in 3D

We may use the multipole powers to estimate the average number of arms possessed by a DLA cluster in 3 dimensions. Figure 6.18 shows the asymptotic value of the multipole powers for 3D DLA clusters. This graph has a peak around P_5.

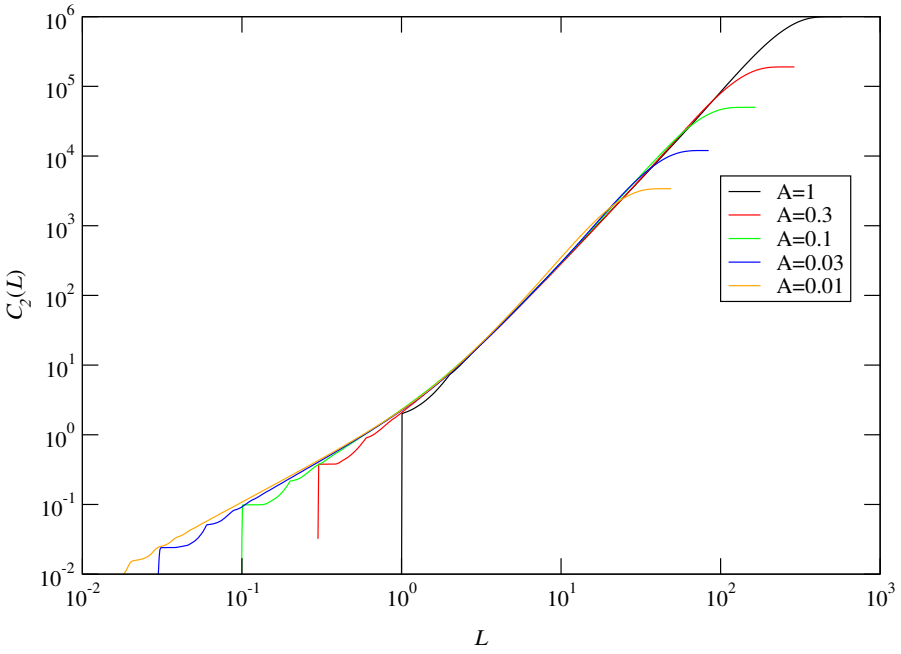

Figure 6.15: The 2 point correlation function for clusters grown in 3 dimensions, times an effective noise reduction, ϵ^{eff}. This noise reduction is chosen so that a data collapse is seen for small L.

Unfortunately, it is not simple to use this to extract the number of arms possessed by a DLA cluster.

Each of the multipole powers are composed of a set of different multipole moments. Each of the moments responds differently to an n armed cluster. Hence, each multipole power does not correspond directly to a given number of arms. Ball [4] estimated that a 2D DLA cluster has 4.9 arms, and thus each arm occupies 73°. Hence, if we consider each arm in 3D to occupy a cone of solid angle 73°, then we will have an estimate of the number of arms of a DLA cluster in 3D. Having created a number of such cones, it becomes apparent that a 3D DLA cluster can have no more than 9 arms.

It is hard to be sure exactly how many arms the 3D DLA cluster has from this measurement of the multipole powers, but it will be between 5 and 9 arms.

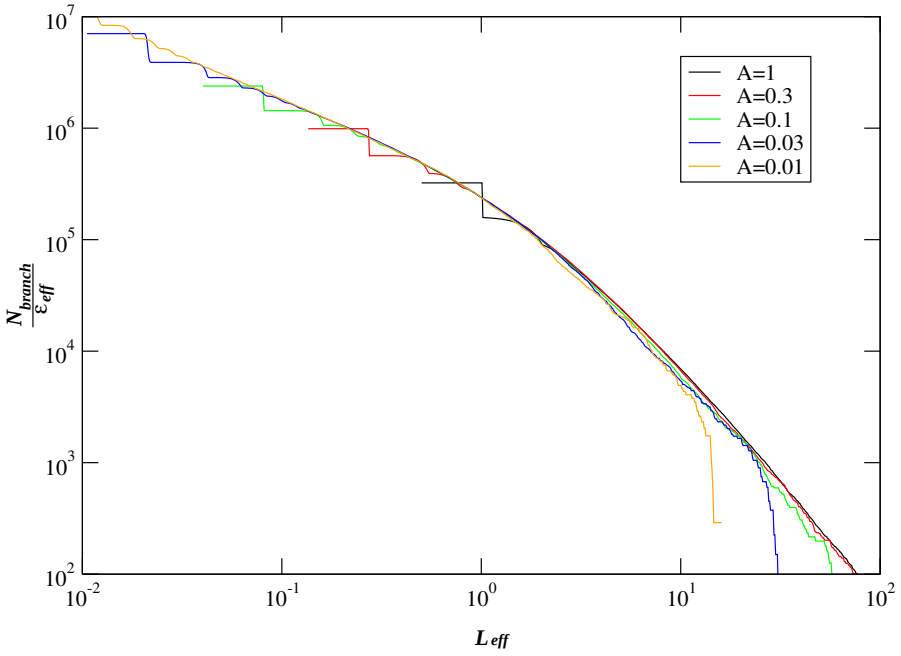

Figure 6.16: The branch distribution divided by an effective noise reduction, $\frac{N_{branch}(L_{eff})}{\epsilon^{eff}}$, against an effective length L_{eff}. ϵ^{eff} and L_{eff} are chosen for $A < 1$ so that a data collapse is seen for small L_{eff}. In general $\epsilon^{eff} < A$, and $L_{eff} < L$ to achieve this collapse.

6.8 Conclusion

Noise reduction off-lattice is a very useful technique for the study of DLA. It has allowed us to elucidate the properties of DLA clusters in 2 and 3 dimensions with greater precision than ever before. This has allowed the renormalisation arguments to be extended to clusters grown off-lattice

The fixed point noise reduction of DLA has been considered. This has been conclusively shown to be $A_f \sim 0.01$ for radial DLA, and not much larger as has been suggested [32]. This has led to the speculation that the geometry in which DLA clusters are grown may have a fundamental effect on their properties. The fixed point noise reduction has been measured in two independent ways and their correspondence has been tested. It was found that both methods agreed for clusters grown in 2D, when the effective noise reduction was taken into account. In 3D, taking into account the effective noise reduction made worse any agreement between the two methods for measuring the fixed point noise reduction.

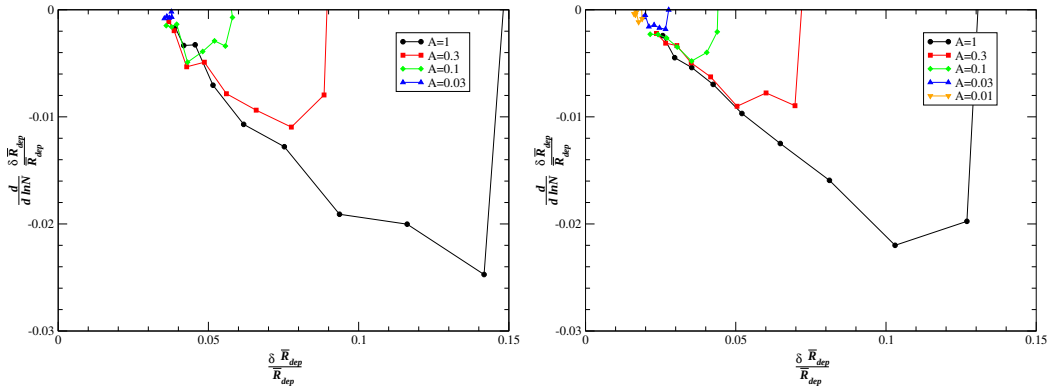

Figure 6.17: The relative variability of R_{dep} in 2D (left) and in 3D (right). Notice that the asymptotic value of the relative variability of R_{dep} is much reduced in 3D, indicating that the bulk of the cluster growth is more stable.

The correction to scaling exponents for 3D DLA clusters was considered. It was found that P_2 shows influence of a significantly faster decaying correction to scaling than those shown by $\frac{\xi}{R_{dep}}$ and P_3. The correction to scaling exponents for higher multipole moments did not show the drift to higher values seen in 2D, and remained roughly constant. The correction to scaling exponent for $\frac{\xi}{R_{dep}}$ is the slowest of all the corrections, suggesting that the argument of Somfai et al. [104] in some sense holds in 3D (see sec. 5.4.2).

The number of arms of a 3D DLA cluster was considered, and it was found to have between 5 and 9 arms. However, the multipole powers are not a good way of measuring this property.

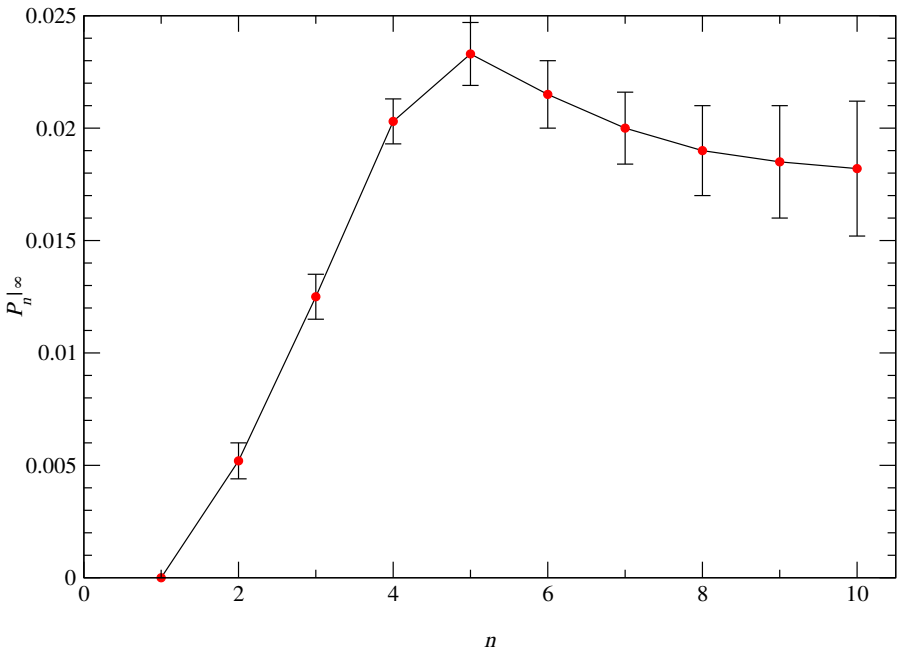

Figure 6.18: The asymptotic value of the multipole powers for clusters grown in 3D. Since the asymptotic value of P_5 is largest, one estimates that the cluster has between 5 and 9 arms.

Chapter 7

Conclusions

We start off confused and end up confused on a higher level. – Alan Chalmers

The role of noise has been examined in two areas - optimisation and diffusion limited aggregation.

7.1 Stochastic annealing

Fink & Ball [34] have introduced a stochastic optimisation method in which the imprecision in measurements provides the analogue of the thermal fluctuations used in simulated annealing. They noted that, in the case of Gaussian noise, the thermal fluctuations may be exactly reproduced by the introduction of an acceptance function.

The exact reproduction of thermal fluctuations has been extended to numerous error distributions. Therefore, stochastic annealing may now be claimed as a generalisation of simulated annealing to stochastic optimisation problems. To arrive at this result the error distributions have been assumed to be a function of a single variable, although this restriction may not in general be necessary. For many of the error distributions considered, the acceptance function was chosen according to $A(x) = 1$ for $x < S_o$. This could not be maintained for the modified Gaussian distribution, and no general criterion is known for when this choice may be made. Hence a more general method is needed which can find rules for such error distributions.

7.2 Probabilistic travelling salesman problem

The probabilistic travelling salesman problem was used as a test-bed for stochastic annealing. Scaling arguments suggest that the optimal tour may be approximated by a sort on scales of less than $1/p$ cities and is TSP-like on longer scales. This leads to the following behaviour of the expected tour lengths

$$E(\bar{L}_{\text{pruned}}) = \beta_{\text{pruned}}(p)\sqrt{np} \tag{7.1}$$

$$E(L_{\text{a priori}}) = \beta_{\text{a priori}}(p)\sqrt{\frac{n}{p}}. \tag{7.2}$$

Simulations of the PTSP using stochastic annealing lead to the conclusion that

$$\beta_{\text{pruned}}(p) = 0.872 - 0.105p \tag{7.3}$$

$$\beta_{\text{a priori}}(p) = \begin{cases} \frac{1}{1.25-0.82\ln p} & \text{for } p > 0.03 \\ \beta_0 & \text{for } p < 0.03. \end{cases} \tag{7.4}$$

One corollary is that solving the PTSP does not optimally solve the TSP on all its sub-problems. The results for the expected a priori tour length need to be confirmed directly. An interesting question is whether the scaling picture can be made more precise, to account for the corrections in $\beta(p)$. This work also naturally extends to the much studied vehicle routing problems.

7.3 Oil field optimisation

The optimisation of oil field projects under uncertain production variables was considered to demonstrate the utility of stochastic annealing in a commercial setting. This posed a number of challenges that are not present in academic problems. The approach of stochastic optimisation was shown to produce significant improvements in the value of the project. The value of the production variables being known more precisely was considered, and this was shown to improve the expected return of the project, as well as reducing the uncertainty in the return.

The introduction of oil price and discount factor uncertainty would be an important development. This would necessitate projects to be treated as multi-stage developments, and may lead to increased returns.

7.4 Noise reduction in DLA

A new technique for noise reduction in off-lattice diffusion limited aggregation has been introduced. The clusters produced possess the same asymptotic properties as "normal" DLA clusters. The measured fractal dimensions are $D_2 = 1.712 \pm 0.003$ and $D_3 = 2.50 \pm 0.01$ for clusters grown in 2 and 3 dimensions, respectively.

The relation between this method and the method of Hastings & Levitov [44] was considered, the numerous advantages of this new method being noted. The HL method has the advantage of ease of measurement of the harmonic measure. However, whether better results for the harmonic measure may be obtained for DLA clusters grown with the fixed point noise reduction remains an open question. The effect of anisotropy on DLA has not been touched on in this work. Much progress has been made on the effect of anisotropy in 2D [31]. Recent work [6] suggests that anisotropy in 3D may be a rich field for new behaviour. This method of noise reduction off-lattice has the potential to be extended to other growth models, with the prospects of new insight in those areas.

7.5 Correction to scaling exponents

Differential plots were introduced for the easy measurement of the asymptotic properties and correction to scaling exponents of DLA clusters. The asymptotic value of the relative penetration depth, $\frac{\xi}{R_{dep}}$, was found to be ~ 0.125 in 2 and 3 dimensions.

In 2 dimensions, DLA is consistent with all quantities being influenced by a single slowest correction to scaling exponent, with the higher multipole powers showing influence of faster decaying corrections. These powers may also show the slowest correction to scaling as well. This suggestion is reinforced since the exponents of the powers do not show continuous drift if they are normalised differently (eq. 5.24).

In 3 dimensions the correction to scaling exponent of P_2 is significantly different from the exponents for P_3 and $\frac{\xi}{R_{dep}}$. The exponents for the higher multipole powers do not show the drift seen in 2D, further suggesting that the powers in 2D will also show influence of the slowest correction to scaling.

The multipole powers allow us to probe the number of arms of a DLA cluster.

It was found that clusters have $\sim$ 4 arms in 2 dimensions and $\leq$ 9 arms in 3 dimensions. The multipole powers, however, are not the best way of measuring this.

Account will need to be taken of the possibility of DLA having more than one correction to scaling exponent, and hence a more complex renormalisation scheme may be necessary.

7.6 Renormalisation of DLA

The renormalisation of the noise reduction has been considered as this is seen to be the key parameter controlling the growth. The noise reduction of the fixed point was measured in 2 different ways - the asymptotic value of ϵ^* (eq. 6.8) and the fastest noise reduction, A_f, to attain the asymptotic properties. The two ways of measuring the fixed point noise reduction are in reasonable agreement both in 2 and 3 dimensions, the value being small, $\epsilon^*_{2D} \simeq 0.003$ in 2D and $\epsilon^*_{3D} \simeq 0.006$ in 3D.

It was noted that clusters grown for small values of A are more compact than one might expect. This has led to the consideration of an effective noise reduction, ϵ^{eff}, associated with each value of A. This improves the correspondence between ϵ^* and A_f in 2 dimensions, but makes worse any agreement in 3 dimensions. The extremal radius of the cluster may not be the correct radius to use in the evaluation of ϵ.

Little work has been performed in cylindrical geometry, and whether the two geometries are intrinsically different remains a question.

Bibliography

[1] M. H. Alrefaei and S. Andradóttir. A simulated annealing algorithm with constant temperature for discrete stochastic optimization. *Management Science*, 45 No. 5:748–764, 1999.

[2] S. Andradóttir. A global search method for discrete stochastic optimization. *SIAM Journal on Optimization*, 6 No. 2:513–530, 1996.

[3] T. Aukrust, M. A. Novotny, D. A. Browne, and K. Kaski. Asymptotic behavior and noise reduction in diffusion-limited aggregation models. *Physical Review A*, 39 No. 5:2587–2592, 1989.

[4] R. C. Ball. Diffusion limited aggregation and its response to anisotropy. *Physica A*, 140:62–69, 1985.

[5] R. C. Ball and R. M. Brady. Large scale lattice effect in diffusion-limited aggregation. *Journal of Physics A*, 18:L809–L813, 1985.

[6] T. Barford and J. Moulding. The anisotropy effects of cubic lattices on diffusion limited aggregation in 3 dimensions, 2000. Final year project, University of Warwick.

[7] P. W. Barker and R. C. Ball. Real-space renormalization of diffusion-limited aggregation. *Physical Review A*, 42:6289–6292, 1990.

[8] J. J. Bartholdi and L. K. Blatzman. An o(n log n) planar travelling salesman heuristic based on spacefilling curves. *Operations Research Lett.*, 1:121–125, 1982.

[9] M. T. Batchelor and B. I. Henry. Growth and form of zero-noise diffusion-limited-aggregation on the cubic lattice. *Physica A*, 233:905–918, 1996.

[10] E. M. Beale. On minimizing a convex function subject to linear inequalities. *Journal of the Royal Statistical Society*, 17B:173–184, 1955.

[11] J. Beardwood, J. H. Halton, and J. M. Hammersley. The shortest path through many points. *Proceedings of the Cambridge Philosophical Society*, 55:299–327, 1959.

[12] A. Benveniste, M. Métivier, and P. Priouret. *Adaptive algorithms and stochastic approximation.* Springer-Verlag, New York, 1990.

[13] D. J. Bertsimas. *Probabilistic combinatorial optimization problems.* PhD thesis, M.I.T., 1988.

[14] D. J. Bertsimas, P. Chervi, and M. Peterson. Computational approaches to stochastic vehicle routing problems. *Transportation science*, 29 No. 4:342–352, 1995.

[15] D. J. Bertsimas and L. H. Howell. Further results on the probabilistic travelling salesman problem. *Eur. J. of Operational Research*, 65:68–95, 1993.

[16] D. J. Bertsimas, P. Jaillet, and A. R. Odoni. A priori optimization. *Operations Research*, 38 No. 6:1019–1033, 1990.

[17] R. M. Brady and R. C. Ball. Fractal growth of copper electrodeposits. *Nature*, 309:225–229, 1984.

[18] A. A. Bulgak and J. L. Sanders. Integrating a modified simulated annealing algorithm with the simulation of a manufacturing system to optimize buffer sizes in automatic assembly systems. In *Proceedings of the 1988 Winter Simulation Conference*, pages 684–690, PO Box 1331, Piscataway, NJ, 1988. IEEE.

[19] R. Cafiero, L. Pietronero, and A. Vespignani. Persistence of screening and self-criticality in the scale invariant dynamics of diffusion limited aggregation. *Physical Review Letters*, 70:3939–3942, 1993.

[20] M. Christoph and K. H. Hoffmann. Scaling behaviour of optimal simulated annealing schedules. *J. Phys. A*, 26 No. 13:3267–3277, 1993.

[21] W. J. Cook, W. H. Cunningham, W. R. Pulleyblank, and A. Schrijver. *Combinatorial optimization.* John Wiley & Sons, Chichester, 1998.

[22] S. P. Coy, B. L. Golden, and E. A. Wasil. A computational study of smoothing heuristics for the travelling salesman problem. *European Journal of Operational Research*, 124 No. 1:15–27, 2000.

[23] G. B. Dantzig. Linear programming under uncertainty. *Management Science*, 1:197–206, 1955.

[24] B. Davidovitch, M. J. Feigenbaum, H. G. E. Hentschel, and I. Procaccia. Conformal dynamics of fractal growth patterns without randomness. *Physical Review E*, 62 No. 2:1706–1715, 2000.

[25] B. Davidovitch, H. G. E. Hentschel, Z. Olami, I. Procaccia, L. M. Sander, and E. Somfai. Diffusion limited aggregation and iterated conformal maps. *Physical Review E*, 59:1368–1378, 1999.

[26] B. Davidovitch, A. Levermann, and I. Procaccia. Convergent calculation of the asymptotic dimension of diffusion-limited aggregates: scaling and renormalisation of small clusters. *Physical Review E*, 62 No. 5:R5919–R5922, 2000.

[27] B. Davidovitch and I. Procaccia. Conformal theory of the dimensions of diffusion-limited aggregates. *Europhysics Letters*, 48 No. 5:547–553, 1999.

[28] B. Davidovitch and I. Procaccia. Dimension of fractal growth patters as a dynamical exponent. *Physical Review Letters*, 85 No. 17:3608–3611, 2000.

[29] L. P. Devroye. The uniform convergence of nearest neighbor regression function estimators and their application to optimization. *IEEE Transactions on Information Theory*, 24:142–151, 1978.

[30] A. K. Dixit and R. S. Pindyck. *Investment undeer uncertainty.* Princeton University press, 1994.

[31] J. P. Eckmann, P. Meakin, I. Procaccia, and R. Zeitak. Growth and form of noise-reduced diffusion-limited aggregation. *Physical Review A*, 39 No. 6:3185–3195, 1989.

[32] A. Erzan, L. Pietronero, and A. Vespignani. The fixed-scale transformation approach to fractal growth. *Reviews of Modern Physics*, 67 No. 3:545–604, 1995.

[33] J. Feder. *Fractals.* Plenum Press, New York and London, 1988.

[34] T. M. Fink and R. C. Ball. Optimisation of a randomly distributed cost function. Submitted to Science.

[35] H. Fujikawa and M. Matsushita. fractal growth of bacillus-subtilis on agar plates. *Journal of the Physical Society of Japan*, 58 No. 11:3875–3878, 1989.

[36] R. Gagne and H. Kroger. Geometrical properties of 3d fractal aggregates. *Chaos, Solitons and Fractals*, 7 No. 1:125–136, 1996.

[37] S. B. Gelfand and S. K. Mitter. Simulated annealing with noisy or imprecise measurements. *J. Optimization Theory and Applications*, 62:49–62, 1989.

[38] S. Geman and D. Geman. Stochastic relaxation, gibbs distributions and the bayesian restoration of images. *IEEE Proc. pattern analysis and machine intelligence (PAMI)*, 1984:721–741, 1984.

[39] F. Glover. Heuristics for integer programming using surrogate constraints. *Decision Sciences*, 8:156–166, 1977.

[40] F. Glover. Tabu search - part i. *ORSA Journal of Computing*, 1 No. 3:190–206, 1989.

[41] W. B. Gong, Y. C. Ho, and W. Zhai. Stochastic comparison algorithm for discrete optimization with estimation. In *Proceedings of the 31st IEEE conference on decision and control*, pages 795–802, PO Box 1331, Piscataway, NJ, 1992. IEEE.

[42] W. J. Gutjahr and G. C. Pflug. Simulated annealing for noisy cost functions. *Journal of global optimization*, 8:1–13, 1996.

[43] J. Haddock and J. Mittenthal. Simulation optimization using simulated annealing. *Computers and Industrial Engineering*, 22 No. 4:387–395, 1992.

[44] M. B. Hastings and L. S. Levitov. Laplacian growth as one-dimensional turbulence. *Physica D*, 116:244–252, 1998.

[45] J. L. Higle. *Stochastic decomposition: A statistical method for large scale stochastic linear programs.* Kluwer Academic Publishers, London, 1996.

[46] J. D. Jackson. *Classical Electrodynamics (second edition).* John Wiley & Sons, Chichester, 1975.

[47] P. Jaillet. *Probabilistic travelling salesman problems.* PhD thesis, M.I.T., 1985.

[48] P. Jaillet. A-priori solution of a travelling salesman problem in which a random subset of the customers are visited. *Operations research*, 36:929–936, 1988.

[49] M. H. Jensen, A. Levermann, J. Mathiesen, B. Davidovitch, and I. Procaccia. Thermodynamic formalism of the harmonic measure of diffusion limited aggregates: phase transition and converged $f(\alpha)$, 2001. cond-mat:0107024.

[50] T. W. Jonsbraten. Oil field optimization under price uncertainty. *Journal of the operational research society*, 49:811–818, 1998.

[51] D. A. Kessler, Z. Olami, J. Oz, I. Procaccia, E. Somfai, and L. M. Sander. Diffusion-limited aggregation and viscous fingering in a wedge: Evidence for a critical angle. *Physcial Review E*, 57 No. 6:6913–6916, 1998.

[52] T. M. Al Khamis, M. A. Ahmed, and V. K. Tuan. Simulated annealing for discrete optimization with estimation. *European Journal of Operational Research*, 116 No. 3:530–544, 1999.

[53] S. Kirkpatrick, C. D. Gelatt, and M. P. Vecchi. Optimization by simulated annealing. *Science*, 220:671–680, 1983.

[54] M. Kolb, R. Botet, and R. Jullien. Scaling of kinetically growing clusters. *Physical Review Letters*, 51 No. 13:1123–1126, 1983.

[55] H. J. Kushner. Asymptotic global behavior for stochastic approximation and diffusions with slow decreasing noise effects: Global minimization via monte carlo. *SIAM Journal on Applied Mathematics*, 47:169–185, 1987.

[56] H. J. Kushner and F. J. Vásquez. Stochastic approximation algorithms for systems over an infinite horizon. *SIAM Journal on Control and Optimization*, 34:712–756, 1996.

[57] J. S. Langer. Instabilities and pattern formation in crystal growth. *Reviews of Modern Physics*, 52 No. 1:1–28, 1980.

[58] G. Laporte, F. V. Louveaux, and H. Mercure. A-priori optimization of the probabilistic travelling salesman problem. *Operations research*, 42 No. 3:543–549, 1994.

[59] P. L'Ecuyer and G. Yin. Budget-dependent convergence rate for stochastic approximation. *SIAM Journal on Optimization*, 8 No. 1:217–297, 1998.

[60] J. Lee and M. Y. Choi. Optimization by multicanonical annealing and the travelling salesman problem. *Phys. Rev. E*, 50:R651–R654, 1994.

[61] I. Lerche. Maximization techniques for oilfield development profits. *Energy Exploration and Exploitation*, 17 No. 6:607–621, 1999.

[62] K. G. Libbrecht. http://snowcrystals.net.

[63] S. Lin. Computer solutions of the travelling salesman problem. *Bell Systems Technological Journal*, 44:2245–2269, 1965.

[64] S. Lin and Kernighan. An efficient heuristic algorithm for the travelling salesman problem. *Operations Research*, 21:498–516, 1973.

[65] Z. Racz M. Plischke. Active zone of growing clusters - diffusion-limited aggregation and the eden model. *Physical Review Letters*, 53:415–418, 1984.

[66] J. A. MacKay and I. Lerche. Maximizing probable oil field profit: uncertainties on well spacing. *Energy Exploration and Exploitation*, 15 No. 3:217–238, 1997.

[67] B. B. Mandelbrot. *The fractal geometry of nature.* W. H. Freeman, New York, 1982.

[68] B. B. Mandelbrot. Plane dla is not self-similar - is it a fractal that becomes increasingly compact as it grows. *Physica A*, 191:95–107, 1992.

[69] B. B. Mandelbrot, H. Kaufman, A. Vespignani, I. Yekutieli, and C. H. Lam. Deviations from self-similarity in plan dla and the infinite drift scenario. *Europhysics Letters*, 29 No. 8:599–604, 1995.

[70] B. B. Mandelbrot, A. Vespignani, and H. Kaufman. Crosscus analysis of large radia dla: departures from self-similarity and lacunarity effects. *Europhysics Letters*, 32 No. 3:199–204, 1995.

[71] M. Matsushita, M. Sano, Y. Hayakawa, H. Honjo, and Y. Sawada. Fractal structures of zinc metal leaves grown by electrodeposition. *Physical Review Letters*, 53 No. 3:286–289, 1984.

[72] McGraw-Hill, London. *McGraw-Hill Encyclopedia of science and technology*, 7th edition, 1992.

[73] P. Meakin. Diffusion-controlled cluster formation in 2-6-dimensional space. *Physical Review A*, 27 No. 3:1495–1507, 1983.

[74] P. Meakin. Diffusion-controlled cluster formation in two, three, and four dimensions. *Physical Review A*, 27 No. 1:604–607, 1983.

[75] P. Meakin. Formation of fractal clusters and networks by irreversible diffusion-limited aggregation. *Physical Review Letters*, 51 No. 13:1119–1122, 1983.

[76] P. Meakin and L. M. Sander. Active zone of growing clusters - diffusion-limited aggregation and the eden model - comment. *Physical Review Letters*, 54 No. 18:2053, 1985.

[77] P. Meakin and S. Tolman. Diffusion-limited aggregation. In *Fractals in the natural sciences*, page 136, New Jersey, 1990. Princeton.

[78] B. Meister, J. M. C. Clark, and N. Shah. Optimisation of oilfield exploitation under uncertainty. *Computers and chemical engineering*, 20:S1251–S1256, 1996.

[79] P. M. Morse and H. Feshbach. *Methods of theoretical physics*, pages 978–980. McGraw-Hill, London, 1953.

[80] W. W. Mullins and R. F. Sekerka. Morphological stability of a particle growing by diffusion or heat flow. *Journal of Applied Physics*, 34:323–329, 1963.

[81] L. Niemeyer, L. Pietronero, and H. J. Wiesmann. Fractal dimension of dielectric breakdown. *Physical Review Letters*, 52 No. 12:1033–1036, 1984.

[82] J. Nittmann, G. Daccord, and H. E. Stanley. Fractal growth of viscous fingers - quantitative characterization of a fluid instability phenomenon. *Nature*, 314:141–144, 1985.

[83] J. Nittmann and H. E. Stanley. Tip splitting without interfacial-tension and dendritic growth-patterns arising from molecular anisotropy. *Nature*, 321:663–668, 1986.

[84] J. Nittmann and H. E. Stanley. Nondeterministic approach to anisotropic growth-patterns with continuously tunable morphology - the fractal properties of some real snowflakes. *Journal of Physics A*, 20 No. 17:L1185–L1191, 1987.

[85] Y. Nourani and B. Andresen. A comparison of simulated annealing cooling stategies. *J. Phys A*, 31:8373–8385, 1998.

[86] J. D. Nulton and P. Salamon. Statistical mechanics of combinatorial optimization. *Phys. Rev. A*, 37 No. 4:1351–1356, 1988.

[87] P. Ossadnik. Multiscaling analysis and width of the active zone of large off-lattice dla. *Physica A*, 195:319–323, 1993.

[88] J. L. Paddock, D. R. Siegel, and J. L. Smith. Option valuation of claims on real assets: the case of offshore petroleum leases. *Quarterly journal of economics*, 103 No. 3:479–508, 1988.

[89] M. Piccioni, R. Cafiero, and A. Vespignani. Monte carlo fixed scale transformation for nonlocal fractal growth models. *Physical Review E*, 55 No. 1:1170–1173, 1997.

[90] T. Rage. Fortran code for the simulation of dla in n dimensions, originally developed by s. tolman.

[91] S. Rees and R. C. Ball. Criteria for an optimum simulated annealing schedule for problems of the travelling salesman type. *J. Phys. A*, 20:1239–1249, 1987.

[92] H. Robbins and S. Munro. A stochastic approximation method. *The annals of mathematical statistics*, 22:400–407, 1951.

[93] S. M. Ross. *Introduction to stochastic dynamic programming.* Academic Press, New York, 1983.

[94] F. A. Rossi and I. Gavioli. Aspects of heuristic methods on the "probabilistic travelling salesman problem". In G. Andreatta, F. Mason, and P. Serafini, editors, *Advanced school on stochastics in combinatorial optimization*, pages 214–227, Singapore, 1987. World Scientific.

[95] P. Salamon, J. D. Nulton, J. R. Harland, J. Pedersen, G. Ruppiener, and L. Liao. Simulated annealing with constant thermodynamic speed. *Computer Physics Communications*, 49:423–428, 1988.

[96] L. M. Sander. Simulations using hl method (private communication).

[97] L. M. Sander. Diffusion-limited aggregation: a kinetic critical phenomenon? *Contemporary Physics*, 41 No. 4:203–218, 2000.

[98] L. M. Sander, Z. M. Cheng, and R. Richter. Diffusion-limited aggregation in three dimensions. *Physical Review B*, 28 No. 11:6394–6396, 1983.

[99] J. Schneider, I. Morgenstern, and J. M. Singer. Bouncing towards the optimum: improving the results of monte carlo optimization algorithms. *Phys. Rev. E*, 58 No. 4:5085–5095, 1998.

[100] S. Schwarzer, S. Havlin, and H. E. Stanley. Multifractal scaling of 3d diffusion-limited aggregation. *Physica A*, 191:117–122, 1992.

[101] S. Schwarzer, M. Wolf, S. Havlin, P. Meakin, and H. E. Stanley. Multifractal spectrum of off-lattice three-dimensional diffusion-limited aggregation. *Physcial Review A*, 46 No. 6:R3016–R3019, 1992.

[102] R. K. Smith. (private communication).

[103] E. Somfai. Simulations of noise reduced dla in 2d (private communication).

[104] E. Somfai, L. M. Sander, and R. C. Ball. Scaling and crossovers in diffusion limited aggregation. *Physical Review Letters*, 83 No. 26:5523–5526, 1999.

[105] J. M. Steele. Subadditive euclidean functionals and non-linear growth in geometric probability. *Annals of Probability*, 9:365–376, 1981.

[106] M. G. Stepanov and L. S. Levitov. Conformal models of laplacian growth with low noise and anisotropy. *Physical Review E*, 63:061102, 2000.

[107] D. N. Sutherland and I Goodarz-nia. Floc simulation: the effect of collision sequence. *Chemical Engineering Science*, 26:2071–2085, 1971.

[108] R. H. Swendsen and J. S. Wang. Nonuniversal critical-dynamics in monte-carlo simulations. *Physical Review Letters*, 58 No. 2:86–88, 1987.

[109] C. Tang. Diffusion-limited aggregation and the saffman-taylor problem. *Physical Review A*, 31 No. 3:1977–1979, 1985.

[110] H. Thome, M. Rabaud, V. Hakim, and Y. Couder. The saffman-taylor instability - from the linear to the circular geometry. *Physics of Fluids A*, 1 No. 2:224–240, 1989.

[111] S. Tolman and P. Meakin. Off-lattice and hypercubic-lattice models for diffusion-limited aggregation in dimensionalities 2-8. *Physical Review A*, 40 No. 1:428–437, 1989.

[112] A. Vespignani and L. Pietronero. Fixed scale transformation applied to diffusion limited aggregation and dielectric breakdown model in 3-dimensions. *Physica A*, 173 No. 1-2:1–21, 1991.

[113] T. Vicsek. Pattern formation in diffusion-limited aggregation. *Physical Review Letters*, 53 No. 24:2281–2284, 1984.

[114] T. A. Witten and L. M. Sander. Diffusion-limited aggregation, a kinetic critical phenomenon. *Physical Review Letters*, 47:1400–1403, 1981.

[115] T. A. Witten and L. M. Sander. Diffusion-limited aggregation. *Physical Review B*, 27 No. 9:5686–5697, 1983.

[116] U. Wolff. Collective monte carlo updating for spin systems. *Physcial Review Letters*, 62 No. 4:361–364, 1989.

[117] S. Yakowitz and E. Lugosi. Random search in the presence of noise, with application to machine learning. *SIAM Journal on Scientific and Statistical Computing*, 11:702–712, 1990.

[118] D. Yan and H. Mukai. Stochastic discrete optimization. *SIAM journal on control and optimization*, 30 No. 3:594–612, 1992.